DOCUMENT RESUME

ED 086 258	IR 000 114
AUTHOR	Swift, D. F.; And Others
TITLE	A Case Study in Indexing and Classification in the Sociology of Education. Development of Ideas Concerning the Organisation of Material for Literature Searching. Report for the Period September 1970-June 1973. Volume One.
INSTITUTION	Open Univ., Walton, Bletchley, Bucks (England).
PUB DATE	73
NOTE	463p.; See also Volume Two, IR 000 115
EDRS PRICE	MF-$0.65 HC-$16.45
DESCRIPTORS	Abstracting; Case Studies; Classification; *Conceptual Schemes; *Educational Sociology; Indexes (Locaters); *Indexing; *Information Processing; Information Science; Information Services; Social Sciences; Subject Index Terms; Technical Reports
IDENTIFIERS	Articulated Subject Indexes; PRECIS

ABSTRACT

The objective of this project was to build a non-interfering indexing system, in which the social scientist users, who have varying conceptual orientations toward their field, would not be restricted in their literature searching by a single conceptual orientation imposed on the system by information scientists. With the cooperation of subject experts, a bibliography and index of a sample of material was compiled and is presented in Volume II of this report. The present volume begins with an introduction and overview which form a preface to a thematically grouped selection of papers. These papers were prepared at different periods in the project, and illustrate the various stages in the thinking and experimentation carried out by the research team. Topics include: 1) early organization of the project, 2) experiments in the appropriateness of the PRECIS indexing system, 3) development of an alternative rationale for indexing, 4) procedures for the selection, processing and description of documents, 5) overall intellectual organization of the index, principles of indexing and the establishment of subject headings. (SL)

ED 086258

A CASE STUDY IN INDEXING AND CLASSIFICATION
IN THE SOCIOLOGY OF EDUCATION

Development of ideas concerning the organisation of material for literature searching

D.F.Swift, V.A.Winn and D.A.Bramer
(assisted from Oct. 1972 by C.J.R.Mills)

Report for the period September 1970 - June 1973

Volume 1

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

The Open University,
Milton Keynes.

IR 000 114

Acknowledgments

We wish to take this opportunity to offer our sincere thanks to Alan Macgregor of OSTI, who has administered our project. We would like to express our warmest gratitude to Mrs Rosemary Campbell for the efficiency and care with which, amongst other things, she has organised the production of this report. We also wish to express our warmest appreciation of all the support Mrs Margaret Malloy has given to the secretarial side of the project since its initiation. Finally, our debt to the subject experts who have cooperated with us in our work is immeasurable, as we have tried to convey in our introduction to this report.

Abstract

This report traces the development of ideas from September 1970 to June 1973 in a project concerned with problems of indexing and classification of social science literature, with particular reference to the sociology of education. An introduction and overview form a preface to a thematically grouped selection of papers. These papers were prepared at different points in the project, and illustrate the various stages in the thinking and experimentation carried out by a research team in cooperation with a group of subject experts.

Contents

Vol. 1

I Introduction 7

II Overview: Development of ideas in course of project 11

III Selected papers

1. Early analyses of problems 37
  - (i) Outline of thinking at September 1970 39
  - (ii) Content and approach of work beginning September, 1970 47
  - (iii) OSTI/Sociology of Education Working Party: Meeting 2 November, 1970. Discussion Paper 51

2. Experiments in indexing 57
  - (i) Experimentation with the PRECIS indexing system, January-July 1971 59
  - (ii) Experimentation with a PRECIS-compatible system, July-September 1971 85
  - (iii) [Experimental system:] Draft guidelines for 'abstractors' and indexers [September 1971] 87

3. Development of alternative rationale 95
  - (i) Methodological issues and information research. 1972 97
  - (ii) Sociology and information science. Paper prepared by Working Party of subject experts: draft as at Easter 1973 183
  - (iii) Responses to our thinking by subject experts and information scientists: correspondence, June-December 1972 205
  - (iv) Outline of current viewpoints, Easter 1973 213

4. Selection procedures 217
  - (i) Journal selection: exercise carried out end 1970/early 1971 219
  - (ii) Item selection: development of procedures January-June 1971 227
  - (iii) Progress report on development of selection procedures, June 1973 235

5. Representation of individual documents 239
  - (i) Development of our ideas about processing individual documents, November 1970 to date 241
  - (ii) Symbolic logic as an aid to preparation of descriptions, 1971 247
  - (iii) Experiment with automatic production of descriptions, March 1972 249
  - (iv) The treatment of material relevant to the sociological analysis of education by existing information services, December 1972 261
  - (v) Working notes on preparation of worksheets, Easter 1973 269

III 5. (vi) Preparation of descriptions: progress report, December 1972 281

(vii) Authors' comments on descriptions, 1971 and 1973 301

6. Principles of overall intellectual organisation 311

A. Work employing a subject heading approach

(i) Suggested principles for reference structure building, May 1971 313

(ii) Handling of single cluster of concepts (sample of material broadly related to achievement): report to Working Party. OSTI/SEA - 29.10.71. 317

(iii) [Overall structure of index.] Proposed principles for indexing: background paper. OSTI/SEA - 19.1.72. 343

(iv) Establishment of subject headings: further considerations of principle. OSTI/SEA - 19.1.72. 359

(v) Report on meeting of Working Party held at Birkbeck College on 19th January 1972 385

(vi) Establishment of subject headings: continued experiment with alternative principles. OSTI/SEA - 24/25.3.72. 387

B. Work employing the notion of dimensions of knowledge

(i) Report on Furzedown College meeting of Working Party, 24/25th March 1972 389

(ii) Report on meeting of Working Party at Birkbeck College, 28th April 1972 395

(iii) Intellectual organisation for material relevant to the sociological analysis of education: aid to literature searching, August 1972 401

(iv) Procedures in constructing a scheme for the intellectual organisation of documents in terms of the dimensions by which knowledge comes to be structured, December 1972 445

(v) Working papers associated with practical exercises in analysis of documents in terms of dimensions, October 1972-January 1973 455

IV General summary, June 1973

V Conclusions

Vol. 2

Sample bibliography and index

I Introduction

This report describes the development of the project, following a preliminary investigation completed in August 1970, up to the end of the current funding period, June 1973. It incorporates the content of several reports submitted to OSTI in the interim. The objective of research from September 1970 has been the development of indexing techniques appropriate to the field of the sociology of education. Whilst associated with the Sociology of Education Abstracts service, the research was in no way limited to the present and particular concerns of the service. The means have been the compilation of a bibliography and index of a sample of material relevant to the practice of the sociology of education. The research team has worked in cooperation with a group of subject experts.

In the first part of the period, the researchers investigated the appropriateness of the PRECIS indexing system for the purpose, and explored ways of overcoming various problems encountered. This experience suggested that an alternative rationale for indexing was required.

The period from January 1972 to date has been devoted to the development of a rationale more in line with the concerns of users searching the literature of the sociology of education, and to preliminary exploration of the most appropriate way of operationalising the ideas which emerged.

This meant a departure from established patterns of research in information science. The task involved the formulation of a theoretical framework as the means for clarifying the nature of suitable indexing techniques, rather than the application of an existing theoretical framework and of techniques defined a priori by such a framework. An implementation/testing research design had therefore to be rejected, and no alternative terms of reference could be formally agreed.

The research team defined the task as one of theory building and operationalisation of ideas, which would provide specifications to guide implementation (the building of an indexing system) and testing (of the ideas embodied in the system relative to the assumptions of the theory from which the ideas are derived) at some future stage. OSTI requested that the work be completed in 18 months rather than the three years proposed. Since implementation and testing phases were dependent on the working through of the ideas to be implemented and tested, condensation of three years work into 18 months has necessarily meant that priority has had to be given to the ideas.

At OSTI's suggestion, the implementation phase will be represented by interim processing, of a select number of the items to be included in our bibliography, in terms of a prototype indexing system which is being developed. By illustrating the current stage of development of this prototype, we aim to provide understanding both of the nature of the system to be developed and also of the method by which we are developing it.

The sample index is one of a series being circulated to subject experts, whose comments will provide the basis for later versions of the prototype. The testing phase, if it can properly be so-called, is thus represented by on-going consultation with subject experts, in particular authors whose work is included in the sample of material with which we are experimenting and the subject experts who abstract and index the material for Sociology of Education Abstracts. Both groups are, in another capacity, users. Responses to our earlier work are reported, but the collection of comments on the current version of our prototype will not be completed within the term of the present project.

The structure and content of this report should be viewed against this background. It comprises a collection of papers thematically ordered and set in context by a chronological survey of our work.

The first volume is prefaced by an overview of the development of thinking in the project over the period in question, and will serve to illustrate the general procedure of the research (thinking at any point in time suggested ideas to be followed up, empirical investigation suggested modification of our thinking and raised further ideas for empirical investigation, in an interplay between theory and practice). More detailed discussion and illustrations of general points will be found in the following chapters, which deal thematically with particular investigations (empirical and theoretical) and illustrate the stage of our thinking at different points in time. Each should thus be viewed in the context of the prior overview. The conclusions which complete the first part of the report are intended to pull together the threads of the various investigations, in relation to the overall ideas, in a way relevant to the practical tasks of system building. The second volume comprises a sample bibliography and index, illustrative of the ideas explicated in the first, and indicative of the nature of the system it is hoped to develop. Detailed data from specific investigations are not included in the report, but will be made available on request, subject to permission from OSTI.

The work which is described here owes immeasureably to the participation of a working party of subject experts, who have met regularly over the entire period covered by the report. We would like to put on record our warmest appreciation of the insights brought to the problems we have studied together, and of the unflagging interest and commitment of the following:

Mr P. Gamage, University of Bristol, School of Education.
Mr M.C.Grayshon, University of Nottingham, School of Education.
Mr S. Hockey, University of Newcastle, School of Education.
Mr R.K.Jones, Staff Tutor, The Open University.
Mr D.J.Oldman, University of Aberdeen, Department of Sociology.
Mr L.E.Watson, Sheffield Polytechnic, Unit for Management in Public Services.

II Overview: Development of ideas in course of project

Origin of our thinking

The origin of our present thinking lies in a seminar with subject experts held at Easter 1970. This seminar had been preceded by questionnairing of British Sociological Association members citing the sociology of education as one of their special interests, librarians serving sociologists of education, Sociology of Education Abstracts abstractors and other groups (eg college of education lecturers in the sociology of education, subject experts using an experimental bibliographical enquiry service run by the project). A range of views concerning appropriate information services to sociologists of education had become apparent.

It would have been possible to make arbitrary decisions as to the form of service we would seek to develop, and there would have been data to back almost any decision we might have made, as is so often the case with user studies. A guiding principle of the project from the outset, however, was that the development of information services should be determined by considerations of the direction of development of thinking in the subject field, rather than the preference of particular segments of users.

The purpose of the seminar, therefore, was to have subject experts weigh the arguments for adopting one form of service rather than another, in the light of their knowledge of thinking in the field, leaving aside their own personal preferences. Whilst it might seem impossible to rule out personal preference, members of the group were quick to tax one another on this score. Consequently, the intellectual issues emerged fairly clearly. A wide variety of sample abstracts, indexes etc was provided to ensure that important issues were not overlooked. Members were asked to think particularly in terms of the printed page index.

The central point which became clear was that differences in conceptual orientation to subject matter were regarded as an important factor in literature searching. Information scientists would not be helping subject experts by adopting one conceptual orientation rather than another, nor by seeking to rationalise the situation. The perception of relationships, and the comparison of alternative perceptions, is something only the subject expert can do for himself. A service which by any means excluded differences in perceptions would be unacceptably constricting, particularly as it was pointed out that it is often impossible to specify search requirements in advance - they rather emerged as the end product of a search.

We have discovered that information scientists find this hard to accept. It certainly brings into question a fundamental principle on which they work, namely that effective information service depends upon subject experts searching the literature in terms of an agreed orientation. Many subject experts deny that an agreed orientation is either possible or even desirable. In our field no single conceptual orientation commands anything like a majority, so that to select one as basis for information processing is to make a service unuseful, and probably unusable, to the majority of users. Our subsequent work in analysing educational documents has confirmed our then suspicion that educationalists differ as widely in their conceptual orientations as do sociologists working in the sociology of education.

We must reject the argument that for literature searching purposes, as distinct from the practice of the sociology of education, this does not matter. The fact that the requirements of searchers are often likely to be expressed in terms of differences amongst orientations, and that the service will have ignored these differences in its processing of documents, cannot be disregarded. Even if we believe that such differences should be reconciled, we have to accept that at the present time they exist, and that we must fit our services to our users rather than assuming that they will change to suit our principles.

There is, therefore, an issue of value which confronted us at this stage. Subject experts' comments indicated that information science, in so far as it attempted to impose structure on their field, reflected a misconception of their purpose in literature searching. A prevailing information science view may be summed up as recognising that social scientists are different from other scientists, but seeing no good reason why this should be so. This is not an issue of fact, and neither viewpoint can be disproved as such. We took the position that, despite the long tradition of scientific discovery on which information scientists call to support their views, a scientific approach to information science calls for a healthy scepticism. On a practical level, if we did not seek to fit our service to the cognitive behaviour of our users as it is, irrespective of what it ought or ought not to be, our service would not be used. We therefore decided to try to develop what we later came to call a non-interfering system.

Information scientists have tended to regard such an attempt as a negation of information science, but we would ask the reader who holds conventional information science views to suspend his disbelief about the wisdom of the attempt. This can be justified only in practice. At this stage, we are concerned to explain the reasoning by which we arrived at our position, and to follow it through in terms of implications for practice. Consequently, relevant criticism of our arguments will take the form 'this point is invalid because it does not accord with the aim of a non-interfering system', rather than 'this point is invalid because the aim of a non-interfering system is inappropriate'. In other words, we would hope that our ideas would be judged in their own terms, not on the basis of a priori assumptions about what is intellectually, practically or economically viable.

We did not at this stage see any conflict between this aim and the conventional practices of information science. Nor did we interpret this aim literally. Clearly total absence of structure results in a non-system rather than a system. Our expectation was rather that, by applying existing practices with sensitivity to the concerns of subject experts, it would be possible to create a literature searching system catholic in its selection of material, in which a loose or minimum amount of structure of an appropriate kind, providing a range of entry points, would be combined with discriminating descriptions (index entries or abstracts) so as to afford maximum opportunity for the user to relate documents to his particular conceptual orientation. This was the general picture of an appropriate system to emerge from the seminar on the basis of its assessment of the collective needs of the field. Our next step was to explore ways of realising such a system. This task was carried out in co-operation with a small working party of subject experts.

The working party contains both sociologists and educationalists. Their terms of reference are similar to those of the seminar members, namely to comment on different procedures with which we experiment from the viewpoint of non-interference with the literature searching of sociologists of education. In this task they assist us with their collective knowledge of current trends in the field so as gradually to improve our understanding of what constitutes non-interference. We regard criticism of their non-representativeness as a set of individuals as true but misplaced. Our attention is directed to the representation of developments in thinking in our field and, from this point of view, the representativeness of our Working Party lies in the fact that members contribute an understanding of developments over the field as a whole rather than 'standing for' a particular set of interests.

In terms of selection of material, the field can be defined in a number of ways. It is possible but not helpful to restrict a definition of the boundaries to sociological analyses of educational situations (ie studies explicitly employing sociological concepts or models). However the borderlines between sociological and other approaches (eg anthropological, economic) are hard to find, and all situations are in some sense educational ones. It can hardly be denied that all sociology is relevant since the sociology of education employs general theories to afford insights into educational situations. Similarly all educational writing is relevant as raw data for the sociologist of education. Practical limitations require, however, that a selection be made. Taking as our universe both sociology, broadly defined to include selected areas of related disciplines, and education, also broadly defined to include for instance child-rearing in the family, we therefore developed criteria for exclusion of least relevant material.

With regard to the recording of individual documents within the system, we felt that a distinction between an abstract and an index entry, even a one word heading, is a false one. Both represent a summarisation of the content of a document, or some aspect of it, and must do if they are to be of any value in literature searching. A system of codes may be used instead as labels or handles for retrieval purposes. Nevertheless we find it impossible to accept the argument that, whereas abstracts must accurately reflect the content of a document, index headings or classification symbols may serve merely to locate documents in a system and may be regarded as a matter of convenience (ie knowing where the indexer has located a document) rather than as conveying meaning.

We have therefore attempted to devise procedures for preparing document descriptions which characterise documents in miniature (ie excluding least relevant detail) and contain all the cues to which a user would respond in judging the relevance of a particular document had he the original in his hands, and by which he might seek to retrieve it. Points stressed at the seminar guided us. It was apparent that whilst impressionistic description was valuable in an ancillary way, formal descriptions (reflecting the structure of ideas in the original) were considered as essential in recording most types of work. It also seemed that there was an irreducible minimum of information which included conceptual and methodological information as well as substantive content. A number of abstracts were studied and in general members wished to see the same information put more concisely rather than reduced in variety. It was considered that the author's own terms should be retained in the description.

In the case of both the selection and description processes, our practices have been refined and modified in the light of different phases of experimentation. However, in retrospect it is interesting to note that we have found no reason to question the appropriateness of the general principles which emerged from discussion at the 1970 seminar.

In our ideas about the nature of a scheme of overall intellectual organisation, our original interpretation of our blueprint has changed more radically. There is nothing in our selection and description procedures, we believe, that is out of keeping with conventional information science practice, though there is much that might be considered too detailed. Our expectation was that an eventual scheme for intellectual organisation would be similar in nature to existing schemes, but less rigid and less highly structured so as not to prejudice any conceptual approach too greatly at the expense of another. What it took us some time to recognise is that, irrespective of the intellectual content of modern analytic approaches to indexing and classification, there is a conceptual orientation built into the fundamental principles underlying the mechanics of schemes employing such approaches which defeats all attempts to sensitise a scheme to a range of alternative approaches. This conclusion was forced upon us when, after experimentation with the PRECIS indexing system, we sought to develop a scheme to our own requirements but compatible with PRECIS machine codes, and were unable to achieve this.

Phase of experimentation with PRECIS indexing system

We came to experiment with PRECIS following the Easter 1970 seminar. The types of indexes which were illustrated to provide a basis for discussion comprised a postcoordinate index, a rotated faceted index and two other precoordinate indexes - an articulated subject index and a PRECIS index. The two latter indexes ran neck and neck for popularity. This confirmed the expectations we held at that time - that precoordinate indexes were more appropriate, for a printed page tool at least, in a specialised field such as ours. The rotated faceted index was considered by members to usurp the subject expert's prerogative in perception of relationships. But cross-referencing in an alphabetical index is no less open to the same criticism; the rank ordering of indexes clearly represents only a relative judgement. Nevertheless the alphabetic format was preferred and we assumed that, by loosening and diversifying the structure, either an ASI or a PRECIS index could be adapted to our requirements. The ASI was less restrictive than PRECIS in this respect, but we were advised that further development of the system was unlikely, whereas the PRECIS system was currently under development. PRECIS was to be used both nationally and internationally, which offered obvious advantages for cooperation with other services. It was therefore decided to investigate ways in which the PRECIS system might, in the course of its development, come to meet the requirements we had formulated.

In hindsight again, the conclusion we eventually reached was foreshadowed in our study of the PRECIS system even before the seminar. We refer in our August 1970 report to its emphasis on the concrete. More specific problems, however, distracted our attention from this aspect of the system when we came to experiment with it. It was found that our 'irreducible minimum content' includes different levels of relationships (relationships at both theoretical and empirical levels of investigation, and between the two) which cannot be kept distinct. Linguistic complexity (multi-phrase concepts such as 'perceptions of the role of language in socialisation') is liable to give rise to ambiguity as the document description or concept string in which such concepts are embedded is manipulated to generate different entry points. Terminology in descriptions has to be standardised to achieve collocation of related items, since entry points are generated automatically from concept strings.

This procedure can be bypassed by means of a KWOC device, but a KWOC index itself is quicker and simpler to use if this device has frequently to be employed. With the exception of occasions where the KWOC device is used (when filing is on entry point alone), filing is on entry point plus all that precedes it in the concept string. Consequently, since reports of research reflect a tendency to avoid duplication, and particularly in the case of entry points derived from the end of the concept string, items are not grouped together by the system with suitable subgrouping, but form sets of one. Finally, there is no reason to suppose that our users proceed in searching as assumed by PRECIS, by matching' concepts one by one to their requirements; it seems more probable that they will often wish to compare one set of ideas, as a set, with another, from the point of view of possible insights into some research problem.

Our experience was that, with care, it was possible to optimise concept strings to represent irreducible content with no more than an acceptable level of distortion and ambiguity but, in this case, the need to retain authors terms and at the same time to provide suitably broad entry points required frequent resort to the KWOC device. Where concept strings are prepared with the generation of suitable subject headings foremost in mind, they usually fall far below the irreducible minimum in content, as well as replacing authors' terms by others. We could find no way within the system of optimising both representations of individual documents and level at which subject headings are stated, at the same time. It is further doubtful that our subject experts would have accepted the system package for generating structure amongst headings, given their view that the perception of relationships is something only the user can do for himself. Tree structures were being developed intuitively by the PRECIS team at the time of this exercise. However, this was at too early a stage of development for us to evaluate.

All this is merely to say that our needs were not reconcilable with PRECIS policy. The system is geared to specific entry and to representations of documents coextensive with the document. The problem for us, it seemed, lay not in that our descriptions were too complex for the PRECIS machinery to handle, but in the pegging of level of subject headings to the level at which individual documents are represented and in the need to choose between authors' language and controlled language as the medium for both. Some features of the system were very attractive to us, in particular the presentation of terms in context, and the facility to modify verbal links to convey the same conceptual relationship even when manipulation of the concept string causes the linguistic structure of the string to be altered. Compatibility with the PRECIS system would clearly be desirable for an information service, given widespread adoption of the system. Our next step, therefore, was to explore the possibility of devising an alternative but PRECIS-like system which both met our requirements and could be processed by PRECIS machine codes.

Experimentation with a PRECIS-compatible system

This proved technically feasible. The problem of distinguishing between different levels of relationships was resolved by operating at the level of the major relationships in a description, and preserving 'chunks' containing minor relationships intact in the concept string. The approach with entry points was to 'round up' authors' terms to a level of greater generality, and to make exclusive use of the KWOC device. Further refinements within the lines of this general approach were judged by computer experts to be viable. Problems in gaining access to computer facilities might have arisen had we gone further along these lines. However, we did not have to face them for, irrespective of its technical aspects, subject experts considered the approach inappropriate on intellectual grounds.

We accepted initially the PRECIS notion of presenting concepts in context, We built into our descriptions a distinction between substantive content and conceptual model as contexts. This was considered entirely appropriate, together with procedures for manipulating the description to ensure that entry points were juxtaposed with the segment of the description to which they related (ie presented in relation to context). Objections were raised when we extrapolated the notion of context to relationships amongst subject headings in an effort to overcome a problem encountered previously in linking our subject headings.

We had had difficulty in creating a reference structure among headings generated from PRECIS concept strings. Part of the problem was that headings often represented fragments of concepts (eg social class extracted from consciousness of social class). But even where the concept itself formed the heading the effect was to bring together material employing a given term for ideas of somewhat different kinds. We first distinguished between use of a term as a descriptive term, as a variable and as a construct, and noted that a different kind of reference structure seemed to be required in each case. Tree structures seemed appropriate in the first case, whereas variables tend to be linked rather in terms of formal relationships and constructs tend to form loose clusters. It was assumed that, with broader headings, it would be possible to employ a clustering technique throughout and to employ substructures relating either to formal or substantive aspects of the ideas associated with given clusters. Further possible formal subdivisions later identified included definitional (contrast between alternative definitions of a concept) and reflexive (perceptions of ability as contrasted with ability) ones. A distinction was made between major, minor and contextual variables. Substantive subdivisions we experimented with included defining characteristics of sample studied, other variables or constructs investigated, and situational ones (eg educational/occupational achievement) were also considered. Indexes exemplifying various combinations of options were examined by subject experts.

Whilst this clarified our ideas, however, it took us even further away from effective cross-referencing among main headings. In the first place, it was found that depending on which mode of subdivision was employed, so the boundaries of a cluster shifted. Secondly, it was impossible to create mutually exclusive clusters in a manner which was generally meaningful. Boundaries which were sociologically meaningful cut across boundaries meaningful to the educationalist, and vice versa, in such a way that distributed clusters could not be brought together by referencing because the principles on which clusters were defined were incompatible, and hence distributed clusters were not systematically distributed but completely dispersed.

Underlying this experimentation was the notion of multiple frameworks of subject headings, defined in terms of context of use, as contrasted with the single framework defined in terms of content of terms. Our expectation that systematic definition and linkage of the headings could be achieved at the level of the individual framework, and that modes of subdivision could be selected for appropriateness to context, proved ill-founded. Our first ideas had been to distinguish between 'real world' and theoretical approaches to subjects of investigation (comparable to that made in our descriptions) as one which would be generally meaningful to our users. A methodological framework was also proposed. Raised from the level of the individual document to that of the corpus, sorting material on this basis was seen to result in a confounding of ideas, apart from creating a problem of bulk in that many documents contained elements of both. However, it might be that the basic notion of framework was sound, but that the particular definition employed was inappropriate. A number of different operationalisations of the same basic distinction (eg units/structure-process) were found to have the same effect. Other distinctions (eg tool building/using, context specific or not, level of analysis) were introduced, and we experimented with them singly and in combination. In these cases, agreed definitions were possible though rather fine lines had to be drawn. However, the product was deemed to be still inhospitable in terms of the range of users' conceptual orientations. It is now possible to see that, because we were thinking and talking about the problems in the terms of conventional information science, we were trying to identify the phenomena (constructs or observables) featuring in our documents, and expecting to find bases for structure in the inherent nature of these objects. This necessarily biassed us towards a rigid normative view of our subject matter and excluded other conceptual orientations from consideration. However, the nature of our problem only gradually became clear.

Phase of rethinking

It was the intractibility of this problem which revived our earlier reservations about the PRECIS emphasis on the concrete. We had previously regarded this as a tendency, in analysing subjects, to focus on substantive aspects of subjects at the expense of theoretical aspects of authors' work, and to admit of a limited range of conceptual models. We believed that the system did not preclude an alternative emphasis. We now began to see that this might reflect a more fundamental assumption about the nature of the subject matter dealt with in documents. PRECIS exemplifies the analytico-synthetic method, which involves breaking subjects down into their basic elements and putting the elements back together according to a prescribed conceptual scheme. Such a method must necessarily involve an assumption of concreteness in the sense of a belief that elements of a subject can appropriately be handled as if they have meaning independent of each other, and of the subject as a whole. It was use of this kind of method that was proving such a stumbling block for us. Indeed, far from solving our problem of creating a system which allowed the user to respond in terms of his own conceptual orientation, we had compounded it by proposing separate frameworks of headings.

Other major approaches within information science were equally unacceptable. Natural language systems handle words in documents as objects in just the same way that analytico-synthetic systems handle elements of subjects as objects. A Dewey-type system which treated sociology and education as separate subject fields would be not merely practically unhelpful but conceptually inappropriate.

There were two lines of thought which might lead us out of our dilemma. First, it might be that we had not gone far enough along the lines of multiple frameworks of subject headings to make an impact on our problem. Secondly, it might be not our conception of framework, but the conception of 'the subject heading' which was inappropriate. This second possibility would call either for new ideas or the conclusion that our material could not effectively be organised for retrieval.

The notion of frameworks was further explored. The working party proposed that a distinction be made between tradition (eg 'isms', great men), area of investigation (to what theoretical or substantive body of enquiry is the author seeking to contribute), and 'real world' (details of sample, physical location of study, and measures taken). Each document was to be classified within as many frameworks as was relevant. The first and third would depend upon explicit statement on the part of the author, the second would allow of reasoned flair. No cross-referencing between frameworks would be required since each defined the total investigation reported in a document from a quite different viewpoint or, in other words, represented a different 'face of knowledge'. Although the different relationships established between documents and scheme resulted in a product deemed to be rather closer to the goal of a non-interfering system, there was still some confounding in the conceptions of the different frameworks which made the treatment of a number of documents problematic. For instance, some authors define the tradition in which they are working in terms of the phenomenon under investigation. Additionally some material (eg prescription and description) fell outside the scheme, since it simply did not seek to afford knowledge of the kind to which the frameworks referred.

We returned therefore to first principles. The problem we had encountered with all the approaches we explored (either established approaches or experimental approaches modelled in some way on established ones) was that none afforded more than a partial fit with our documents. What we had encountered was not merely a question of overlapping sets amongst which we had to draw boundaries and map interrelationships. Even when we employed several discrete frameworks there was some material which was not classifiable within the scheme. Conceptual complexity was not the only source of our trouble. Even in the current three frameworks scheme, we could not divide our material into even three mutually exclusive categories, at least on the basis of evidence in the documents, with or without license for reasoned flair.

The very existence of traditions, professional organisations etc within the sociology of education is evidence of communication within the field, even if there is a polarisation in terms of discipline affiliation where professional identity is concerned. The increasing call upon sociologists to carry out investigations which guide educational policy (the Plowden Report is an obvious example) indicates a measure of communication between sociologists and educationalists. The persistence of these phenomena suggests that there is something here upon which an information service may build in organising the literature of the field. Social and professional contact, however, does not mean that there are necessarily shared understandings. It is the cognitive organisation of the field in which we are interested. If we agree that a measure of cognitive organisation exists, such that members of the field are able to locate themselves in groups, why then have we not found it in conceptual analysis of the literature of the field where it must surely be at least partly manifest.

We were forced to consider the possibility that the problem lay in assuming that the literature in a given field comprises a universe of knowledge which the information scientist may map by analysing the subject matter to which documents refer in terms of the elements they comprise, and the relationships amongst them, and by defining this as the universe of public knowledge. Whether a scheme provides for analysis of content or meaning in terms of concepts, words, word fragments or any other element, and whether the bits are put back together ab initio by the system designer, or at the search stage by an information officer or user, the same implicit assumption seems to underlie all the schemes. The assumption is the objectivist one that meanings are to be regarded as entities which remain constant when divorced from context.

In fields such as chemistry, perhaps, the term for a chemical compound may have a constant relationship with that compound, and alternative definitions may not be in question. Analyses of documents in terms of substructure can thus validly be made as if it were the compound itself being analysed. Situations involving social phenomena are not so clear cut, and documents about them cannot be analysed in this wholly objectivist way. Not only do terms reflect layers of meaning, some of which remain implicit, but documents must be treated with recognition of prior considerations raised by the investigation of social phenomena.

Social science investigations are guided by differing models of man. Some social scientists certainly view man as a passive object acted on by the social environment and, in the case of documents reporting their investigations, the information scientist may appropriately adopt an objectivist mode of analysis. Other social scientists take a range of different viewpoints. For instance, some regard the individual as constructing his social environment in interaction with significant others. In such a case, use of objectivist analysis by the information scientist would lead the user to interpret the document quite differently from its author, even to focus on different entities, and hence to distort its meaning and confound user relevance judgements. It seemed to us, therefore, that the particular mode of analysis assumed by current techniques of indexing and classification was too restrictive for the range of viewpoints represented by our material.

One would have expected that, by presenting terms embedded in their 'context of use', a system such as PRECIS would have a built-in safeguard against such restrictiveness. In the first place, however, the most general context recognised by PRECIS is the geographical, and the next one which is in current use is the key system or the system acted upon (the equivalent of the dependent variable), and this tends to be defined substantively following the train set by statement of location (eg university, teachers). Additionally, the system preference for 'passive order' (eg Country. System A. Role of/in. System B.) biases it towards a particular (linear-causal/path analytic) kind of orientation. Secondly, and more importantly (since we could see nothing in theory to prevent rethinking of this aspect of the system to increase its hospitality to a wider range of conceptual models), the derivation of subject headings from descriptions of the contents of documents means that the overall organisation of material is one which takes no account of alternative models of man. Again, in theory, there seems to be no reason why the algorithm could not be extended by this further intellectual step. We came to the conclusion that the problem lay in a purely literal conception of the notion of context, as one of words conjointly describing the content of a document, which is not alone sufficient as a basis on which reliably to impute meaning. Despite the potential we believe the PRECIS algorithm has to represent documents in relation to the theoretical framework or conceptual model employed (which could be exploited to the full if concept strings did not have to be written with the requirements of generating subject headings from them in mind), the algorithm is such as to present all conceptualisationfrom a single objectivist viewpoint and must necessarily distort much of that material.

Reflections on these considerations suggested that, if our reasoning forced us to question the basic assumptions upon which existing systems are based, the only way forward lay in proposing alternative assumptions and exploring the implications for information processing. This was not a step which we took lightly. However, consideration of the nature of the existing research which supports assumptions underlying existing retrieval systems encouraged us to believe that it was a justifiable step. Existing systems were developed to a large extent, where the processing of specialised material is concerned, on the basis of experience in serving natural and physical scientists. These fields tend to exhibit conflicting conceptual orientations at the extreme research front but, once accepted, conceptual orientations tend to be long-lived by comparison with those in the social sciences, and to share a common objectivist standpoint. Systems for social scientists seem to be modelled on those for 'hard scientists'. Studies of the information needs of social scientist users, like our own earlier ones, tend to offer choice from given alternatives rather than to assess the appropriateness of the range of alternatives.

Additionally, internalisation of the orientation assumed by current indexing systems may be expected to reduce system designers' and operators' capacity to 'see' a document in any other way. This has a bearing also on the appropriateness of the criteria of evaluation which are employed in investigating the effectiveness of their systems, and which users become conditioned to regard as valid. In other words, it seems probable that the situation operates to legitimise current practices. However, to the extent that we were investigating such practices in relation to a more general criterion (non-interference) external to them, our conclusions thus far, whilst unexpected, must always have been a logically possible outcome of the line our investigation had taken. The appropriateness of non-interference as a criterion, if not the interpretation we had placed upon it, was confirmed rather than discredited by the judgements of our working party, and other subject experts who studied the output of our various empirical exercises, not least by the reasoned nature of the criticisms they addressed to our products. We set ourselves therefore to develop an alternative rationale.

Formulation of an alternative rationale

We have noted the apparently dominant assumption in information science that the meaning of information is independent of the language in which it is expressed, and of the context of its use in terms of the perspectives of author and user. This assumption appears to run right through the process of research and development in information science. We considered the kinds of conceptual model employed in designing information systems. 'Systems' and 'flow' models, for instance, are widely employed. Typically the conceptual model represents a view of the information situation as one of the user as passive recipient of information channelled to him by the information scientist. Even where user opinion or feedback has helped to shape the 'channels', effective information service is held to require fixed channels so that, providing the user can state his requirements, knows the labelling system, and 'pushes the right button' or has an intermediary do it for him, he can obtain the information he needs. Such a model is not necessarily made explicit, but an implicit objectivist viewpoint may reasonably be assumed. Other aspects of information science research tend to support this assumption of objectivism. In regard to method, for instance, we noted that an emphasis on hard data could often lead to data collection guided by a definition of variables determined as much by accessibility or amenability to measurement as by appropriateness to the situation. In the same way, selection of technique of analysis may fail to take account of the nature of the assumptions entailed by the technique (eg that the factors revealed by a factor analysis of a body of data are equivalent to its intellectual dimensions, even though these factors may not be susceptible to a plausible interpretation). We found, too, that even where information scientists attach importance to user needs, these are defined in relation to the aggregate (a sum of individuals) rather than in terms of the collective needs of the group. Social survey and experimental modes of investigation tend to be valued above smaller scale case study and participant observation approaches for the greater reliability of the information they afford, irrespective of the quality of the insights afforded.

Despite surface variety in approaches, there seemed to us to be a single paradigm here, the elements of which collectively take their validity from a central assumption concerning the objective nature of information. We had been forced to question the validity of this central assumption, and because of this the prevailing methodology as a whole came into question. This left us little to build on, and by early 1972 it was clear that it had become necessary to devise an alternative methodology.

This was a major turning point in our project. Our working party of subject experts provided our starting point. They argued that the meaning of terms used to describe the content of documents does not necessarily inhere in the phenomena under investigation, nor in the documents describing them, but may depend upon assumptions external to both. Therefore the framework of an effective indexing or classification scheme must similarly represent factors external to the subject matter with which it deals. In other words, material must be presented in terms of the dimensions of knowledge by which it comes to be structured. However, to the extent that meanings remain partly implicit, both because they have multiple sets of associations (everyday and technical) which cannot always be enumerated, and also because a writer may assume that he can rely upon a certain background of experience amongst those he is addressing, we must assume a measure of user involvement as an integral element in system development, use and maintenance, for which information science skills will be able to devise no alternative. In other words, both organisation and retrieval processes will be situations not of the system acting for the user, but of the user as interacting with and part of the system. In terms of use of the system, we might expect that search strategies would be formulated as much by the questions 'who thinks like me' or 'who thinks like ...' as by 'who has been thinking about ...' or 'whose research produced data about ...'. Users sharing a 'thinks like ...' definition would make similar relevance judgements, but such judgements would differ from group to group, and the overall pattern of judgements would be expected to change over time as the ideas in the field as a whole changed. This conceptualisation of the retrieval situation is of fundamental importance from the viewpoint of the insights it affords into what will constitute a non-interfering retrieval system.

The view of information which emerges here does not assume that knowledge is independent of the knower. At the same time, there is no espousal of a position of epistemological relativism. The assumption is rather one of patterned difference in the imputation of meaning, and of patterned change over time. There are several points to clarify here. Meaning is regarded not as a property of a document or of an individual, but as a property of an interaction between the ideas implicit in a document and the ideas a user brings to the perusal of it. The nature of this interaction is expected to vary systematically according to the orientations of the actors. As such, this is a social situation, and the nature of the interaction is a cognitive one. Since the orientations of the actors are liable to change, we must view this as an ongoing process, rather than as a static situation. These assumptions are incompatible with a view which defines literature searching as a task prior to, or a means towards the intellectual activity of studying the documents retrieved in order to acquire or generate knowledge. On this view, literature searching is itself an intellectual activity involving the same social-cognitive process, even though it involves interaction with condensed representations rather than original documents. This view may be contrasted with a 'technological process' view, which assumes that information can be subjected to processes of manipulation without changing its properties as information. On the social/cognitive view, such manipulation will be liable to disrupt the dynamics of the underlying social/cognitive processes, and the outcome will be misinformation or false knowledge, leading to relevance judgements later found to be inappropriate. A relevant analogy at the individual level, perhaps, is the disruption of the learning process which can result from the design of teaching programmes in isolation from consideration of the developmental sequence of the child, and of the sorts of situation in which the child finds himself. The design of information systems similarly must be such as to function in accord with the social/cognitive characteristics of the social situations within which they are to operate rather than being imposed upon them. An appropriate conceptual model, therefore, would be one which characterised the social/cognitive processes involved in the activity of literature searching, to which our information system would be the hand-maiden, rather than regarding an information system as encapsulated from them.

These ideas both emerged from and were crystallised by a survey of recent sociological literature. In particular a recommendation to study Berger and Luckmann's 'The social construction of reality'[1] shaped our thinking. Berger and Luckmann see subjective experience and objective knowledge as simply different and interdependent ways of viewing the same phenomenon. What starts as subjective experience comes to be viewed as objective knowledge through consensus, as individuals transform their experience into some sort of symbolic representation that others will understand. Ideas are never made fully explicit, and understanding rests also upon a background of shared experience as individuals interact in groups. Different communities develop their own rules for agreeing on what is valid knowledge or, in other words, on what is an appropriate way of looking at the world. Such 'agreements' represent perspectives. Whether we think of knowledge relating to everyday situations or of professional or discipline knowledge, which is rooted in everyday experience but rests also upon 'agreements' particular to the sub-community, new knowledge becomes objectified by framing it in a particular perspective and validating it according to the rules of that perspective. However, this is not a one-way but a dialectic process. Rules for determining what constitutes valid knowledge are pragmatic. The nature of the situations to which objectified knowledge refers is always subject to change. Thus, goodness of fit between knowledge and situations (or social situations at least) necessarily diminishes over time. At a certain point, therefore, the rules of a given perspective are liable to become inappropriate and to be changed, though not necessarily pari passu with those of other perspectives. Knowledge that was formerly regarded, provisionally at least, as objective, then becomes a matter of subjective belief when no longer implicitly maintained as objective by consensus. In the same way, what was regarded as a matter of subjective belief can come to be regarded as objective knowledge. A forthcoming paper by the Working Party incorporates this viewpoint.

Whilst most information scientists would perhaps accept that we should think in terms not of 'public knowledge' but of 'public knowledges', it is the dialectic aspect of the 'knowledge process' which may be hard to accept. Enquiry in the natural and physical sciences is more commonly regarded as having a cumulative 'brick-upon-brick' nature. Kuhn[2], however, has shown that scientific thinking has changed paradigmatically over time. Information scientists sometimes have tended to adopt Kuhn's ideal type of one community/one paradigm and to suggest that the shifting schools of thought in the social sciences are a sign of immaturity. But Kuhn[3] himself has questioned this notion in his recent writings, as a possible oversimplification, if only in that whether a situation is a mono- or multi-paradigmatic one depends on the definition one takes of community. Thus a given field may be equally validly characterised as comprising a number of mono-paradigmatic communities or as representing a single multi-paradigmatic, and hence pre-paradigmatic, community. Whichever view one were to adopt concerning the future development of the social sciences, we felt that Berger and Luckman offered a plausible explanation of the situation currently obtaining. Even were it desirable to accelerate acceptance of a single paradigm in a given field, there is absolutely no means by which to evaluate current paradigms for this purpose other than by social scientific activity such as is currently in progress. Consequently, we were led to the view that there is no choice for information system design but to take as its framework the current broad differences amongst perspectives and to keep all options open.

We combed the theoretical literature further in order to elaborate our model in greater detail but, whilst this exercise generated a number of hypotheses concerning relations between perspectives in the knowledge process, which we documented for future exploration, it did not help with the immediate task of realising the notion of perspective in constructing a system. We considered techniques employed by socio-linguists who stress the need to study culture, from the viewpoint of their potential for the investigation of perspectives. Frake[4], for instance, suggests a need to discern 'how people construe their world of experience from the way they talk about it'. Unfortunately we found that sociolinguistic investigation along these lines is at an early stage and techniques are as yet under development. Techniques of content analysis, citation analysis and similar approaches are relatively well developed, but lack the theoretical underpinning by which their appropriateness for the purpose in hand may be judged. Thus, even given the cooperation of subject experts in interpreting or assessing any analysis we might undertake, we were forced to reject an objective approach to the task in favour of a more intuitive one as more appropriate to a non-objectivist position.

Taken as a whole, this all added up to a methodology of a very different nature to that currently dominant in information science. We were anxious to test the logic of our reasoning, which we worked out in the course of preparing a paper submitted to OSTI in mid-1972, and this paper was circulated to a number both of information scientists and subject experts working in areas related to our own. We had been at pains to state different positions taken within information science, to state grounds for questioning them, and to discuss the arguments relevant to the position we were proposing instead. It was therefore disappointing to find information scientists responding not by counter-argument so much as by reiteration of the positions we had questioned, without reference to our argument. One comment, for instance, was to the effect that our paper merely indicated that social scientists just did not know what they wanted, when the burden of our paper was to explain why it was necessary to keep all options open and not to impose a single viewpoint. A substantial amount of time was spent in following up some of the comments by letter to clarify some of the major points we felt to be at issue, without notable success. Clearly there was a failure in communication for which we must accept some responsibility. At the same time it was hard to escape the feeling that our readers were interpreting our remarks within a framework of certain preconceptions, and thus our problem in communication was not so much one of explaining what a given issue was as one of demonstrating that there was an issue at all.

Four kinds of preconceptions seemed to emerge from this exchange of views. One was the 'technological' view that information systems are like machines which are best 'thumped to make them go', which tends to suggest that thinking about them is unlikely to help in diagnosing factors in poor performance. Secondly, there was the pragmatic view which seeks the best value for money from what is currently available, with the effect that certain goals are ruled out of consideration from the start as, say, too expensive, irrespective of the possibility that it might be more helpful, and cost no more, to go part way towards a superior goal as compared with complete achievement of several inferior ones. Amongst those of a more theoretical orientation, there is a viewpoint which holds that all theory is potentially relevant and we may draw from it without regard to the assumptions implicit in any particular theory - this may be termed an 'eclectic' position. A further viewpoint is one which makes fully explicit the objectivist position we described earlier and contrasted with a dialectic one, with the consequence that individuals espousing these two positions tend to 'talk through each other', to use a Kuhnian phrase. Some of these views are discussed by our working party in their forthcoming paper.

Responses from subject experts took a somewhat different line. There were two main schools of thought. One thought the ideas valid and important, and went on to point to areas of debate which were inadequately argued through, whilst accepting the general line of argument. The other school of thought denied the possibility of constructing an effective indexing system on the grounds that, by definition the construction of such a system involves the imposition of rules, whereas those of this particular school of thought, equally by definition, view a system in terms of the practices of members - these two notions are incompatible. More constructively, apart from a number of specific points, and some general ones concerning presentation of the argument, the call in the main was for further elaboration of the argument, and in particular the conceptualisation of the knowledge process. All the points will be taken account of in revising the paper, but pressure of time prohibits us from following them up in the term of the current project.

Operationalisation of alternative rationale for information processing

Our next step was to explore ways of operationalising our ideas. The notion of perspective is a very slippery one. Bernstein[5] has noted some broad distinctions (order/control, interdependence-dependence/conflict-voluntarism, negotiated encounters/structural relationships, members' categories/observers' categories). Such distinctions in themselves proved too broad for practical purposes. Another set of distinctions which comes immediately to mind is one in terms of 'isms' (eg structural functionalism, Marxism). Definitions, however, tend to vary substantially, and it has been pointed out that perspectives tend to unpack into a number of elements. Wallace[6] suggests that they must be defined in terms of the dimensions they comprise. He proposes a scheme in which major 'isms' are classified primarily by simultaneous reference to the way in which the social world is defined and the way in which it is explained. Wallace contrasts objective with subjective definitions, and explanations imposed on the social world with those generated by it. Pairs of characteristics correspond to four sets of perspectives (further subdivision is provided for). We were able to build on Wallace's ideas to take account also, within a modified schema, of the growing body of phenomenological material. Hoyle[7] classifies social theories of education, in a way similar to Wallace, in terms of nomothetic/idiographic and radical/conservative dimensions. We explored other classifications of perspectives also. The working party's earlier thinking on 'traditions' was relevant here too. In particular the 'great men' notion was extended to that of 'significant other' to characterise perspectives defined in terms of discipleship to particular people. A distinction was made between adoption of theories or models associated with the thinking of a particular individual, investigation of a research question raised by the work of a given individual, and use of a particular method developed or espoused by a given individual.

The working party had also come to posit two basic modes of search in association with the notion of tradition or perspective and this influenced our thinking at this stage. They distinguished between the researcher wishing to locate his already formulated ideas within a tradition (ie tracing a genealogy for his ideas), and the researcher starting with a research problem suggested by a body of data and wishing to identify ideas or methods which might afford relevant insights into his problem or provide a language in which to talk about it. For the latter purpose we drew upon the paradigms proposed by Riley[8]. Riley focusses upon the operational decisions the researcher makes in abstracting the salient features of a situation under investigation in a search for an explanation which fits the situation. We selected, and modified for our purpose, the paradigms relating to the case for study, the universe for study, sociotemporal context and the properties featuring in the investigation. These paradigms might be said to sum to another way of defining perspective.

A sample of material was processed along these lines and circulated towards the end of 1972 amongst subject experts external to our working party. Comments received suggested that this attempt was along the right lines, since the principles we were employing were meaningful to subject experts and there was a basis for reasoned comment. Comments tended to focus on the way in which we had proceeded in order to realise the notion of perspective. One point was that the scheme did not allow of fine enough distinctions amongst perspectives. Yet finer distinctions would require a high degree of sensitivity on the part of the classifier. In the eyes of subject experts, categories defined in this way carry different prestige rankings, and there might be unintended labelling effects. Overall the effect could be counter-productive for a non-interfering system by reifying the notion of perspective.

Even though we saw this as a scheme which subject experts would help to refine, and as one in which subject experts would carry out the classification process, thus allaying fears on the earlier scores, the problem of reification was more serious in that this would impede the development of new ideas. In seeking ways of overcoming this problem, we found a means also to minimise the possibility of bias inherent in such a classification even when implemented by subject experts.

In effect, whereas we had been trying to classify existing perspectives in terms of a priori notions about the dimensions they comprise, we set ourselves instead to identify and classify the dimensions by which subject experts characterise perspectives in talking about them. The elements in such a classification were thus available without restriction and could be used in combination to construct perspectives to fit particular situations. Elements could be added to the scheme without disrupting it, and new perspectives could be generated as required. Such an approach, in relating to perspectives as used, would not, unlike its predecessor, be expected to reinforce tendencies towards reification, although it would reflect such tendencies where they were present in the thinking of subject experts.

We hoped to be able to cull dimensions from general theoretical discussions and overviews of sociology and of education. In fact, in the case of sociology we identified few dimensions additional to those built into the classification schemes we had already encountered. Relevant discussion in the case of education was confined to the writing of philosophers of education and to comparison of different discipline approaches to the study of educational situations. The effect was that a substantial proportion of both sociological and educational literature was unclassifiable in terms of the available dimensions. We therefore carried out further analysis of the research literature itself during the early part of 1973. On the basis of this work, the outline of a scheme has now been established which embodies many of the notions explored earlier, but is wholly open-ended and flexible. Separate frameworks have a facet-like function in distinguishing at the level of the document as a whole between fundamental methodological assumptions, level of analysis, nature of theory or conceptual model, themes of enquiry (theoretical, substantive and population to which a study is generalised), and significant other, together with distinctions (for situations where it is relevant to consider elements in an investigation in isolation from context) between variables, sample and location of study. Parallel sets of frameworks correspond to different purposes in study (understanding as an end in itself/understanding as required by some external purpose). Unlike faceted schemes, a given framework does not necessarily present a set of mutually exclusive categories, but may list a range of alternative dimensions on which material within the framework may be categorised. Frameworks, too, are defined formally rather than in substantive terms.

In principle, such a scheme could be implemented in either post-coordinate or pre-coordinate fashion, though in practice cumbersomeness in use would restrict post-coordinate operation to an on-line computer situation, and the complexity of the description of many documents would detract from a pre-coordinate presentation in either coded or natural language format. For this reason the guiding conception of the purpose of the scheme may be considered more appropriate, namely to provide a preliminary analysis of material of a kind which would allow of the subsequent construction of perspectives relevant in literature searching, leading to a classified presentation of material by perspective, accessed either via systematic or alphabetic lists of dimensions represented in perspectives. Procedures for preparation of descriptions require no drastic modification to include all the data required as input for classification. Such a scheme is able to handle sociological and educational material by the same principles and, in that it is framed in terms of general research considerations is, in conception if not in specific content, equally applicable to other social science fields.

Value in use can clearly be determined only in use. In so far as the active participation of users is concerned both in the project, and in the Sociology of Education Abstracts service, our experience is that subject experts are ready to devote time to tasks that seem to them meaningful and worthwhile and that, on the whole, subject experts are willing to accept, because they know and can make allowances for, the intellectual judgements of other subject experts. We see no reason to doubt that subject experts do want to engage in the kind of literature searching activities they say they do. In practical and economic terms there is little in SEA's experience to suggest that the proposals are not viable in this sense. Objection has been raised to the notion of a reworking of material every ten years or so to take account of changes in views in the field. We do not see the selection of 'established items' of the previous ten years, and their reworking in the system as it stands for publication prior to the beginning of the next ten years, as an impossible task but rather as a process of review which will be welcomed. It is unfortunate that much of our analysis will be irrelevant to other systems in the event of any cooperative exchange, and that reanalysis of material from other systems would be required fully to integrate such material into our system. We believe however that in general the advantages will outweigh the disadvantages.

References

1. Berger, P.L. and Luckmann, T. The social construction of reality: a treatise in the sociology of knowledge. Harmondsworth, Penguin Books, 1967. (First published in U.S.A. in 1966.)
2. Kuhn, T.S. The structure of scientific revolutions. (International Encyclopaedia of Unified Science, Vol.2., No.2) Chicago, University of Chicago Press, 2nd ed., 1970.
3. Kuhn, T.S. Op. cit.
4. Frake, C.O. 'The ethnographic study of cognitive systems', in Fishman, J.A. ed. Readings in the sociology of language. The Hague and Paris, Mouton, 1968.
5. Bernstein, B. Unit 17, in Swift, D.F. et al. School and society course. E.282. Bletchley, Open University, 1971.. (Restricted circulation at time of writing.)
6. Wallace, Walter L. Sociological theory. London, Heinemann, 1969.
7. Hoyle, Eric 'Social theories of education in contemporary Britain', in Social Science Information, 1970, 9(4), 169-186.
8. Riley, Matilda White Sociological research. I A case approach. New York, Harcourt, Brace and World, 1963.

III Selected papers

1. Early analyses of problems

The three papers in this section describe our thinking about the problems of information handling in our field at the beginning of the period covered by this report. The first two papers are retrospective, and were included in a report prepared for OSTI in December 1971, and the third is a working paper which was circulated in November 1970 to the subject experts with whom we were working.

The first paper surveys the findings of our previous work, and offers a formulation of the requirements of an information system for our field, as we understood them at the time.

The second paper describes how, guided by this formulation of requirements, the next phase of work was organised.

The third paper attempted to analyse the problems for information processing, and to raise important issues on which the research team wished to consult the Working Party of subject experts.

III 1. (i) Outline of thinking at September 1970

The present report deals with the second phase of a project, sponsored by OSTI (the Office for Scientific and Technical Information),on information processing and the sociology of education. The second phase is concerned specifically with the exploration, in cooperation with a Working Party of subject specialists, of indexing techniques. The broad guidelines for this work were formulated as a result of a prior phase of work involving detailed study of SEA (the Sociology of Education Abstracts service) and consultation with sociologists of education. The first phase has been written up in detail in a report submitted to OSTI in 1970[1]. It is not intended to duplicate information recorded there in the present report. The purpose of this paper is simply to indicate the stage of our thinking at the start of the second phase of the project.

Relation between project and SEA

Indexing techniques do not function in a vacuum, and the analysis of SEA, apart from leading to specific suggestions regarding the development of the SEA service, provided a picture of the kind of environment within which indexing techniques would be expected to operate. But experimentation with such techniques in an operational situation is liable to produce disruption in an ongoing service, both organisationally and in the eyes of its users. Organisational innovation, beyond a certain point at least, cannot be justified without a high degree of certainty of adequate returns, whilst the returns cannot be assessed unless innovation is made.

It was therefore decided that the project should compile and index a bibliography independently of SEA, though the work would profit by the lessons learnt from SEA, as a means of testing out ideas developed in the course of thinking about SEA.

With regard to the scope of the material to which indexing techniques should in the first instance relate (generalisability to other social science fields was seen as a question for later study), the study of SEA showed that no simple definition in terms of subject content was adequate. It pointed to a number of other criteria of relevance which should perhaps be taken into account, depending on whether the viewpoint is 'what the sociology of education is' or 'what sociologists of education do'. (It should be noted that SEA abstractors receive no editorial guidance as to the selection of items from the journals for which they are individually responsible, and it was found that inter-abstractor reliability was low.)

In terms of characteristics of abstracts, SEA abstracts lack uniformity. (Again SEA abstractors accord such treatment to a document as, in their individual specialist judgment, they consider appropriate.) This may be attributed, in some measure at least, to the differing nature of material and to questions of relevance (we have to consider not only degree but also kind of relevance, in particular direct/indirect relevance). Whilst abstractors may be in accord as to the way they categorise documents, the treatment given to different categories appeared to vary across abstractors. Hence, relative to the document, SEA abstracts were found to lack predictability.

For some purposes this constitutes no disadvantage. As a basis for indexing, however, it was found that extensive reference to the originals was needed to supplement the contents of abstracts, and some notion of an irreducible minimum of essential information began to emerge, applying equally to abstract and index entry as successive condensations of the content of an original. From the point of view of searching either index or abstracts this notion appeared to be important. Essential information should be taken to refer not merely to the substantive topic(s) or variables handled in a given study, but also to other considerations (how these elements are related in the writer's mind or work, the perspective from which the work is viewed, the purpose in undertaking it, and so on).

Insights of these kinds were separated from organisational aspects of a functioning service to help to determine the general directions along which the compilation of a bibliography and the indexing of it should proceed.

Relation between project and sociologists of education as users of information.

A major effort was put into consultation with users in the first phase of the project, and this supplemented insights into SEA thinking in guiding the later work described in this report. An important feature of this effort was that the intention was not that of the conventional user study (the study of users) but was an attempt to study the problems with users and from the point of view of users.

This meant that we were not concerned to relate users to predetermined categories, with the aim of adjusting either users or services to achieve better fit. We were rather concerned to discover the reasons for which users did or did not fit such categories, and to adjust the categories for better fit with users' aims and objective as far as we could ascertain them.

A preliminary survey of members of the British Sociological Association declaring an interest in education and of education librarians[2] provided a broad picture of use, and problems in use, of formal services. It also enabled us to identify a select group expressing interest and/or concern with information seeking problems, who were asked to consider and compare alternative solutions in a more detailed way. This group formed the nucleus at a seminar, held in Oxford at Easter, 1970, in which discussion focussed upon the kind of assistance subject specialists would like to receive, in literature searching, from an information service in the light, on the one hand, of procedures which are technically feasible and, on the other hand, of developments in work in the sociology of education which it is desirable either to support or to discourage.

Present information seeking practices are of interest as indicating the user's adjustment to the characteristics of services currently available, and for what can be inferred about the purposes underlying these practices from the ways in which users fail to conform to information scientists' expectations. The 'ought' questions to which we proceeded provide a basis for determining priorities in terms of the purposes to be catered for, some conflict being inevitable. Certain kinds of information processing techniques were identified as more appropriate than others (eg precoordinate v post coordinate indexing). The precise ways in which our generalised formulation of requirements should operationally be implemented have been the substance of the second phase here reported.

More specifically, it was found that, in general, type of use (eg retrieval/current awareness) was not associated with type of service (organised to aid systematic searching/browsing). The 'concerned' group were characterised by more extensive use of information services but the failure to relate operational use to the nature of the service still obtained.

Case study data collected, independently of surveys and seminar, in connection with both use and non-use of a bibliographical enquiry service run for experimental purposes, supported a belief that, although users may well lack expertise in use of information services, the acquisition of such expertise would do little to aid them in their searches. A more important factor seemed to be inappropriate intellectual basis of services. The relevance judgments to be made either involved factors of which indexing or classification schemes take little or no account (indirect relevance, slant, peripheral/central to topic), or involved definitions of problems which cut right across such schemes.

In sum, in terms of individual needs, a service seemed to be called for which was fairly wide ranging, in which material was briefly characterised and descriptions broadly grouped and thus perhaps unsophisticated, but which allowed of high discrimination in selection according to purpose from material relevant in subject content.

With regard to collective needs, and the 'ought' questions involved, it was accepted (some even insisted on it) that users would have to scan a number of items once they had located potentially relevant documents in a service which, outwardly at least, followed conventional lines. Both the perception of relatedness amongst documents, and evaluation of documents was seen to be the individual expert's job. Even subject experts could not do this for other subject experts, though if subject experts rather than information experts without appropriate subject background were responsible for information processing, more license would be acceptable.

The main point made was that no single framework was generally accepted, this meant that no framework could be regarded as more valid or useful than any other, and that the imposition of any one framework created distortion for those who wished to view material within some other framework. This was seen to apply both to the organisation of a collection of descriptions of documents, and to the content of individual descriptions.

A number of sample indexes and types of abstracts were discussed from the point of view of the way in which subject specialists might be helped or hindered in their searches.

Guidelines for experimentation.

Perhaps because users are necessarily conditioned by what they know, few positive suggestions with regard to information processing techniques emerged either from the broad survey of sociologists of education or from more detailed consultation at the seminar. This was not in fact expected. There were, however, clear views as to what an information service should not do, which were helpful as guidance for experimentation in indexing.

In general it seemed clear that the following requirements would be unlikely to be challenged:

(1) In selecting material for the bibliography we should take a broad rather than a narrow definition of relevance.

(2) It was decided to begin experimentation, though without commitment, with a study of the PRECIS system.[3]

(3) With regard to amount and nature of data to be provided about individual documents, this should constitute a complete though miniature representation, both in order to avoid bias and also to permit of access from a variety of approaches.

(4) Organisation of material should be broad and general rather than specific and tightly structured. It should also be relevant to the kind of enquiries likely to be searched, yet tied to no particular mode of defining (intellectually) these enquiries.

(5) The main kind of conflict to be resolved was that between the approaches and kind of assistance required by those with a sociological background and by those lacking such a background. It was agreed that prior consideration should be given to processing material in ways which sociologically were appropriate. However, to the extent that additional procedures could assist non-sociologists without involving anything the sociologist would regard as distortion, the aim should be to provide such assistance.

(6) A further difference amongst user approaches received consideration. This is the difference between those in which the purpose of enquiry is the understanding or development of theory and those in which the focus is on the understanding or solution of practical problems. Implicitly or explicitly there is considerable overlap here, but in operational searching terms these were felt to deserve separate consideration.

(7) Compatibility with techniques employed in existing operational services should receive a lower priority than effective handling of the material from the points of view of the specific user groups concerned. The grounds were that only once all the

possibilities (within reason) had been explored and considered would it be possible to determine what procedures should be retained at all costs and what might be sacrificed in the interests of compatibility.

It was taken as given that the procedures developed must be associated with clearly formulated principles, and not simply with individual judgments, if they were to achieve any measure of consistency or generalisability. Equally it was recognised that judgment was inevitably involved in realising such aims, both in formulating and/or applying appropriate procedures, and hence that close contacts with subject experts were essential. Assessment by eventual actual use would be expected to result in further development and refinement of techniques as well as confirming the general lines along which the initial statement of requirements had been operationally implemented. Ultimately, in the event of conflict or even alternative compromises, a decision would rest with the Director of the Project, involving a considered judgment as to the best interests overall of the potential clientele.

A final point should perhaps be made here with regard to interpretation of the so-called findings of user studies. However great the care with which data are collected, the researcher necessarily shapes the eventual picture which emerges from the data, simply by the questions he has asked as related to the questions he has not asked. Bias will almost certainly have entered in, in that the kind of responses collected are determined by the way in which questions have been asked. Finally, the actual responses are interpreted within the framework of the researcher's beliefs and expectations of the sort of picture he requires or considers relevant. By consulting respondents not merely in terms of specific questions but also with regard to the general framework within which information problems are viewed by those concerned about them, we hope that such bias as remains is shared with those of our respondents who have some understanding of the nature of the problems, and is thus not likely to detract from the immediate outcome of our work. Appropriateness in the context of groups of experts in other fields is a separate and later question.

Dec. 1971.

References

1. Swift, D.F. et al. *Investigation into Sociology of Education Abstracts: report on first stage of project*. (Report submitted to Office for Scientific and Technical information.) Oxford, 1970.

2. Education librarians reported use of formal services by their readers. To the extent they themselves are users, librarians would not be expected to differ from, say, sociologists in searching associated with sociological enquiries. The sociological enquiry itself would be expected to suggest the logic of the search; librarians may have the advantage in speed in searching, but a mismatch between logic of index and logic of search will create the same problems for librarian and sociologist alike. The extent to which the logic of the search can be adjusted to that of the index without altering the nature and possibly the point of the enquiry is a matter of subject background rather than solely one of technical (index) expertise.

3. This is a system devised by Derek Austin of the British National Bibliography, and is the subject of a current research project headed by Mr. Austin.

III 1. (ii) Content and approach of work beginning September, 1970

Following the principle of consultation with subject experts a meeting was called with two main objects: to seek general advice on implementing the general guidelines outlined on the basis of prior investigation, and to set up an on-going study group or orking arty of subject experts to appraise the effects of alternative procedures tested in the course of experimentation. To give a focus to the investigation of problems in information processing, it had been decided to compile a bibliography of material relevant to the sociological analysis of education.

Invited to the meeting were those members of an earlier seminar who had indicated a wish for continuing participation in our work, together with all Britain-based sociologists, educationalists and others who prepare abstracts for SEA. Around twenty people attended.

They were reminded of the objectives of the project by an information paper circulated beforehand. Draft guidelines for SEA prepared as a by-product of the previous phase of the project, and background information about subject indexing, indicated the general lines along which we were thinking. The Director of the Project, who chaired the meeting, stressed 'the possible dangers of the well-intentioned but often ill-informed application of machine methods of information processing, developed in the context of information services for natural and physical scientists, to the literature of the behavioural sciences'.

With regard to the composition of the Working Party to be set up, the following outline was given:

"It is hoped that there will be a 'hard core' of 8 - 12 members, meeting about every two months. The majority will be sociologists with a special interest in education, but there will be also some educationalists with a special interest in the sociology of education. Some will be SEA abstractors. Information and other subject specialists may be co-opted on the basis of special interests".

The Director stressed that it is not intended that the Working Party should be representative of sociologists of education. It should however be representative of schools of thought within sociology, not necessarily in terms of individuals but in that members will take account of and speak for viewpoints which may not be their own.

The functions of the Working Party were described as follows: "The precise nature of the contribution to be made by the Working Party can only be established after discussion. It may include:

(a) advice on implementing the broad guidelines already agreed,
(b) assessment and comment as work progresses,
(c) consultation and feedback from colleagues outside the Working Party,
(d) detailed study of specific items typifying conceptual problems and issues, and of alternative ways of handling them,
(e) individual contributions on a special interest basis."

A report on this meeting recorded the following points:

"1. Coverage

Alternative types of frameworks to which selection decisions might be related were discussed. Two main kinds of answers to the key question 'relevant to what (or to whom)' (ie the problem of what the sociology of education 'is') were identified: answers at a conceptual level and answers at a practical level. For the purposes of selection it was decided to work at a practical level.

With regard to the application of a quality criterion both exclusivist and inclusivist positions were strongly argued. The difference between the requirements of a retrieval and a current awareness service was also discussed. It was stressed that whilst the eventual bibliography would be a useful by-product of the work, the central concern was to develop effective information handling procedures. The main criteria for selection therefore should be such as to throw up a comprehensive range of the problems for information processing. The question was eventually resolved by setting priorities.'

The parameters identified as important were:

(a) type of document (journals, 'official' publications including series, books, theses)
(b) type of relevance (SEA's four categories)
(c) date (5 year periods working back from 1970)
(d) provenance (UK and USA, other English language, foreign countries).

"2. Selection procedures

It was decided that as many cells in the matrix thus formed would be dealt with as time permitted. As a start material of direct relevance would be identified in about 20 major journals, British and American. It was decided to experiment with a proforma recording such information about a journal article as would enable a judge to make the same selection decisions as he would make on the basis of the articles themselves. Miss Winn agreed to prepare a first draft of such a form for trial on four suitably chosen journals and members agreed to experiment also amongst themselves. After comments have been received on this pilot test it is planned to go on to the further sample of journals, the results of this exercise to be discussed at the next meeting.

"3. Indexing

Types of indexing model were discussed and it was seen that work in this area could only proceed gradually as the essential problems presented by different types of work were clarified."

In the event, the hard core Working Party has had a membership of around six consistently attending members.

Meetings have been held at roughly monthly intervals. We are greatly indebted to members of the Working Party for their interest in and continuing support of the work.

Dec. 1971.

III 1. (iii) OSTI/Sociology of Education Working Party: Meeting 2 November, 1970. Discussion Paper

There is a feeling that the ground to be covered in discussion should be clarified. These notes (an agenda would hardly be appropriate) are intended to serve to structure our discussion. Ideally the Working Party would make its own analysis of the problems we have outlined, but to save time we are presenting the problems as we see them, on the assumption that important questions to which we have not devoted enough attention will emerge in discussion. It would be naive to suppose that we can 'solve' the kinds of problems with which we are involved. However in the 'information' context we are concerned with conceptual issues and it would be equally naive to ignore them. For practical reasons decisions have to be made. These decisions can at the least be made in the light of the considerations important to sociologists and for the right kinds of reasons.

It seems highly desirable that as early as possible we should move from purely general discussion to a consideration of general principles in relation to the documents we have to deal with. It would be most helpful if members could come prepared to illustrate their points with reference to particular items, if possible drawing examples for detailed comment from a body of material which all may study in advance (say SEA, 6(3)).

(1) General aims of bibliography and index

The detailed planning of bibliography and index will be influenced by the kinds of use to be regarded as legitimate. It clearly cannot be expected to provide 'facts', eg details of findings. But it might be expected to enable the user to provide himself with 'answers' to at least some of the following types of questions (the 'legitimacy' of these or others requires discussion):

(a) What documents exist on a given topic (eg social stratification or, more open ended, birth order effects)?

(b) What are the major works on a given topic?

(c) Can I, from the information in the index, identify the 'good' work on a given topic?

(d) Any of (a) - (c) might be narrowed to major etc. work of a given kind.

(e) More specifically a user might ask: Is any one approaching the problem on which I am working in the same way as me?

(f) What factors associated with my problem have been studied and written up?

(f) Is there work similar to mine in methodology or theoretical approach, in an area with which I am not well acquainted?

The kinds of information to be provided will be determined by the kinds of questions agreed, for practical purposes, to be legitimate. For instance, if the user is to be able to identify 'good' empirical work, it might be decided that he ought to want details of population, sample, control of variables etc. Alternatively it might be felt that bibliographical information indicating the origin of the work ought to give sufficient, more reliable etc. information. If he wishes to obtain references on work dealing with factors bearing on a problem he is studying it would be possible to include cross-references to suggest further headings under which he might search. Alternatively it might be considered that it would be undesirable to offer such a facility. We have here a range of 'ought' questions to consider, as to the kind of information pertinent in a given type of situation.

(2) Content of bibliography

We have drawn a distinction in the guidelines for abstractors (copy circulated previously) between four kinds of material, each having a different kind of relevance for the sociologist working in education:

(a) Sociological analysis of education (direct relevance)

(b) Sociological writing other than that relating directly to education (indirect relevance)

(c) Sociographic or educational data or description (raw data)

(d) Social but non-sociological writing about education other than (3) (background and general interest).

Ultimately the amount of material we include will be determined by the amount of time available to us. However it is useful at this stage to ask whether there is any important category overlooked, and also whether the suggestion that degree of selectivity should vary from one category to another is acceptable. We need to discuss the desirability of including books. A problem with regard to journals is whether to cover 'important' journals or 'important' items (the latter representing of course a far wider range of journals).

The major problem is to elaborate criteria of selection for each category of material, taking perhaps as a starting point those sketched out in the guidelines. Examples of 'borderline' cases would seem to be essential for this purpose.

(3) Content of index entries

The type of indexing system to be used is such that a 'complete' statement of the 'subject' of a document is given in an index entry. 'Completeness' and 'subject' are however ambiguous terms. We need to arrive at a working definition of the desirable content of an index entry which will enable sociologists to produce appropriate (in terms of the use to be made of the index) 'subject' statements, and which will also ensure that different sociologists would produce reasonably consistent statements of the 'subject' of any document. It is from analyses by subject specialists that the more formal index entries will be prepared.

An adequate analysis will clearly take account of both the substantive (empirical setting) content and/or the sociological concepts dealt with, depending whether it relates to an empirical, theoretical or a factual or discursive work. However probably no work is ever perceived by different people in exactly the same way; in particular the individual's own theoretical framework may intervene. That there should sometimes be interpretation seems inevitable.

We need to study ways of reducing differences in analyses resulting from such factors. It may for instance be helpful to specify the standpoint from which a work should be viewed - the author's standpoint/an independent standpoint etc. In the case of work with a theoretical basis, is it feasible to attempt a level of statement more generalised than that of the author or any individual (and thus of course not immediately assimilable within the user's own theoretical framework), but indicative of major theoretical orientation (eg structural-functionalist)?

(4) Structure of index entries

It is further desirable that not only should we ensure a consistent relationship between the content of index entries and that of the documents for which they stand proxy, but also that we should devise ways of ensuring that one index entry is consistent with another for a document of similar content. The way in which this is achieved in the type of system to be used is by means of a generalised schema providing a sequence of categories; according to the category to which any part of a 'subject' statement belongs, so its position in the sequence may be ascertained; the sequence of component parts of an index entry follows that of categories within the schema. (See paper on subject indexing previously circulated.)

The important point is that if the schema used does not provide appropriate categories, either the indexer will take the 'nearest' and this, across indexers, is likely to lead to inconsistency, or there will be no guidance at all and the way in which index entries are structured will become largely arbitrary.

An example from this type of indexing system may serve to clarify the problems as well as to illustrate the general principle;

A section of the schema would be as follows -

1 Viewpoint
2 Active system
3 Effect, action
4 Key system
5 Discipline
6 Environment

Thus, the 'subject' of a document might be stated as follows: 'Influence of racial subculture on intellectual development of school children in Hawaii'

giving: (6) Hawaii (4) School children (3) Development, intellectual (3) Influence (2) Racial subculture

from which this and further index entries would be generated:

HAWAII

School children. Intellectual development. Influence of racial subculture

In principle, were the effect of some variable upon racial subculture the subject of some other document, this system would ensure that on this occasion the entry would read:

HAWAII

School children. Racial sunculture. Influence of ....

Thus the different ways in which racial subculture is considered in the two studies would be immediately obvious.

However, the inappropriateness of categories such as 'active system' is clear. Hence the need for us to seek a schema in accord with the ways sociologists think.

(5) Terminology

As another aspect of the problem of achieving clarity and consistency in our subject index we face the problems of terminology. A main object is to group together material on related concepts under the same heading, so that if different terms are used for the same or virtually the same concept it will be desirable to select one and to make a cross-reference from the others. The grounds on which such a decision may be made require careful consideration. It is clearly important that conceptual distinctions should not be obscured. Indeed (since sociological use may sometimes be sloppy and non-sociological use of a given term different from that of the sociologist,eg 'role') it may be necessary to distinguish between concepts for which separate terms do not exist. It is suggested that a list of concepts prepared for a seminar at Easter should be taken as a starting point for either a thesaurus or glossary of terms. The principles which should guide such a task are a matter for discussion. The problems have not as yet been studied in any detail. It may be possible to draw upon authorities such as Gould and Kolb, but there remain conflicts on which, for practical reasons, we should seek to reach agreement of a 'best possible' kind even whilst this will not lead to any immediate resolution of the conflicts.

III Selected papers

2. Experiments in indexing

The papers in this section refer to the period January - September, 1971. The indexing experimentation which they describe was by no means the only activity in this period. It was however, our central concern during this time, and the papers form a relatively self-contained set.

OSTI arranged that we should experiment with the PRECIS system, and the first paper describes our exploration of the facilities afforded by PRECIS, in the light of the requirements we had previously formulated on the basis of consultation with subject experts.

The best results we could produce did not meet these requirements and, whilst reconceptualisation of the system was in progress (there had been a shift from systems thinking to the ideas of structural linguistics), it did not seem to be expected that these or any of our ideas would have any implications for the actual system. However, exchange of information amongst services is clearly an important consideration. We therefore experimented with a PRECIS-compatible system, an outline of which is given in an extract from a report submitted to OSTI concerning our experimentation with the PRECIS system, and in a paper prepared for our Working Party of subject experts.

In the event, the procedures were found still to be tied too closely to a kind of thinking which makes inappropriate assumptions about our users and their literature, and we had some more fundamental rethinking to do.

III 2. (i) Experimentation with the PRECIS indexing system, January - July 1971

Basis for cooperation

An opportunity to cooperate with the PRECIS team occurred in the first instance through the good offices of D.J. Foskett, Librarian of the London University Institute of Education. The extended contact (Jan. - July 1971) with the PRECIS team was arranged by OSTI. We are grateful both to OSTI and to Mr. Wells, Derek Austin and his colleagues at the British National Bibliography (BNB) for this opportunity of seeing PRECIS from the inside. It was a most valuable exercise in helping us to clarify in more detail the requirements of a retrieval system for the needs of our field.

A misunderstanding which existed should be explained by way of introduction to an account of this block of our work. A distinction was made at the outset between PRECIS as implemented in BNB from January 1971 onwards, and PRECIS forming the subject of an OSTI-sponsored project, involving study and experimentation with a view to development and modification of the system in the light of research. Our experiments with PRECIS were undertaken in connection with the latter but at a later stage the possibility of modification to the system was called into question.

If this was the case, since our experimentation was guided by the belief that there was to be rethinking and modification better to accommodate the needs of specialised fields such as ours, our work was clearly illegitimate. Our work is fully reported here because a considerable amount of time was devoted to it (OSTI were kept fully informed, by regular reports, of the nature and progress of the work), and because we were not aware, at the time, of any inappropriateness in our approach vis à vis the purpose of OSTI'S PRECIS project. The comments that follow represent our views concerning the effectiveness of the system for handling our material in the light of the objectives which we formulated; the effectiveness of the system in relation to other material or other objectives is not under discussion.

The main element in our brief concerned distortion, both distortion resulting from intellectual factors at the indexing stage and that attributable to presentation factors affecting use of an index, though clearly there was a call for exercise of judgment as to the -

extent to which any particular kind of distortion would affect location of relevant material or discrimination amongst relevant documents by users.

A preference for broad headings was discovered in consulting with our users. This might seem incompatible with the nature of the indexes produced by the PRECIS system. PRECIS is geared to entry points representing specific elements of subjects, descriptions of subjects being 'coextensive' with documents. PRECIS has first and foremost to handle the monograph material covered by BNB. Our concern is with a specialised field, and hence elements of subjects will be very specific. However, specificity is relative and, regarded simply in terms of the scope of our field, there would not necessarily be an incompatibility with our requirements.

Learning the system

Work began with several intensive training sessions in which the system was demonstrated and explained, with the aid of examples (mostly at a title-like level) from a variety of fields including education. Notes were taken and the worked examples recorded to supplement the introductory guide to the system. The absence of a manual at the time, and of worked examples covering detailed as well as general indexing, was something of a handicap, although the training received was said to be comparable to that given to new BNB staff. At this stage the object was to translate natural language statements into 'strings' coded by means of 'role operators'. Selected items were indexed in parallel,[1] and examples quoted are drawn from this work except where otherwise stated. At this stage the PRECIS team prepared machine input. Machine coding of strings for the computer was self taught later on.

Our aim was not merely to learn the 'rules' but to understand the rationale underlying the system so that, in studying the system product both vis à vis the content of our material and in relation to use factors, we could determine not merely whether certain kinds of modifications in the entries might be considered to improve the product, but also the extent to which system facilities might be exploited to this end.

The general procedures associated with the system were already familiar. The broad question in our minds was the nature of the theoretical basis of the system and, in particular, the underlying assumptions about the intellectual content of the material to be indexed and about users' intellectual approaches to this material. This appeared to us to be a matter of more than academic interest. If the 'rules' were associated simply with practical objectives (eg collocation), then rules along with objectives might be modified operationally without affecting the performance of the system. In other words, the objectives might be redefined and still be achieved efficiently using the same machinery. On the other hand, if the rules were associated with an intellectual structure to be imposed upon the material, then almost certainly to modify this structure even in a minor way would mean rethinking of the total system.

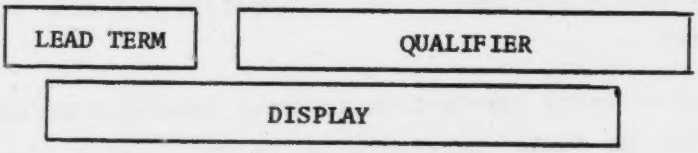
Discussion during initial training sessions repeatedly touched on the issue of learning of rules v understanding of principles. We also began to clarify particular issues for later study; these fall under three broad headings: overall structure of index, physical structure of individual entries, intellectual structure of individual entries.

1. Overall structure of index. The basic feature of the overall structure of a PRECIS index is that the entry points are provided by terms from the document description, which are successively brought into the filing position by means of computer rotation. This ensures that entries in which the same entry point features are juxtaposed (it does not necessarily produce actual clustering, as will be seen later). In principle, descriptions are 'coextensive' with documents, and concepts, as they occur in the documents, serve as entry points. In practice, not only must terminology be controlled, but it is clear that terms serving as entry points must be of a certain order of generality if there is to be clustering. Coextensiveness and this order of generality are largely synonymous when handling monographs (though summarisation is employed in subjects stated in the form 'social aspects of ...'), but this is not necessarily the case with specialised material.

The derivation of entry points from document descriptions was not satisfactory for our field, particularly for the journal literature we were handling. In the first place our concepts were often too specific for this purpose, yet to summarise in ways such as the 'social aspects of ...' form would have meant the loss of major variables. This specificity was associated with linguistic complexity in stating variables, (eg'university's elite perpetuation role', or role of university in perpetuation of an elite) which did not always lend itself to the PRECIS procedures for term extraction from compound concepts. (We return to this point later.) Hence, even juxtaposition of documents in terms of broad areas of interest could not be assured. The construction of a PRECIS reference structure was in its early stages and could not be studied in detail. However, our level of specificity was such that reliance solely on cross-references to bring related material together would have made the index too complicated for ease in use.

This led on to consideration of valid handling of concepts or variables. In one example discussed, 'Universities. Students' was recommended in preference to 'University students' (even if in the given document the students were not studied in the university setting) on grounds of collocation. The distinction which for practical considerations is here allowed to be blurred is perhaps not of great importance. In another example, 'student protest' was said not to be a valid concept. This is much more arguable. In more complex examples author's terminology has usually to be accepted if one is not to run the risk of misrepresenting his ideas, and indiscriminate 'fracturing' of what is a single idea carries the same risk. Degree of fracturability is also closely associated with provision of entry points.

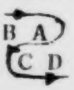
2. Physical structure of document descriptions. Each concept in an entry is said to establish the context of the next, and PRECIS assumes a particular kind of behaviour on the part of the index user, namely that from any chosen access point each entry will be tested, concept by concept, for continuing relevance. To accommodate the expected procedures of use, the description has a two line structure:



Individual concepts or terms required as lead terms are successively brought into position as entry points, ie they become 'lead terms', by the process of rotation as follows:

A BA CBA DCBA
BCD CD D

(Once terms have passed into the qualifier they may be combined into a single phrase or phrases by means of a device called the 'substitute'.)

Even if, intellectually, users proceed in this stepwise fashion, this structure seemed open to question on practical grounds. In the last rotation position, or where there is no display (DCBA), the eye naturally follows the sequence of terms as a sequence (D-C-B-A). In intermediate rotations, the habit of reading horizontally from left to right seems likely to result in the following eye movement: 

(ie reading B A C D), thus disrupting the 'concept in context' notion.

The system is designed so as to avoid such a procedure, and the accompanying risk of misinterpretation by users, both by means of typography and layout of entries. If it occurs, verbal links as well as position serve to preserve relationships amongst terms without ambiguity. The risk of ambiguity in the case of simple subjects is probably small, eg:

Immigrant children. [In] Primary schools
[Their] Educational problems.

With more lengthy subjects, however, there is a 'wrap round' effect, and the eye may be more prone to read on through the qualifier and only then to proceed to the display:

Test Anxiety Scale for Children. Anxiety ratings. Research
into influence of students' academic achievement
on teachers' assessments of personality
compared with Teachers' ratings of personality (24
characteristics) - **Sample population:** 96 children,
54 teachers [2]

The longer the qualifier, the less easy it seems to drop the eye automatically from lead term to display. The case of the lengthy entry was felt to require study.

We felt we should also consider whether the two line structure, even if read as intended, could give rise to any ambiguity or risk of misinterpretation. At this time we had no more than a vague impression that some rotations appeared to be more effective than others and that this might have to do with the relation between the physical structure of the entry and the structure of the subject, as in the following example:[3]

1. Lebanon
   Political development. Perpetuation of elite. Role of American University at Beirut. Influence of students' social class, religion, English language as medium of instruction and high fees
2. Elite. Political development. Lebanon
   Perpetuation of elite. Role of American University at Beirut. Influence of students' social class, religion, English language as medium of instruction and high fees
3. American University at Beirut. Role in perpetuation of elite. Political development. Lebanon
   Influence of students' social class, religion, English language as medium of instruction and high fees
4. Social class. Influence on elite perpetuation role of American University at Beirut in political development. Lebanon
   Students' social class

N.B. A special device is used in order to eliminate additional variables in the display when one variable from a set of variables comes into the lead position (as in 4).

A natural language statement of the subject of the document to which these entries relate might be 'Factors (social class etc) bearing upon the elite perpetuation role of the American University at Beirut in political recruitment and integration in Lebanon'. One problem seems to be that two ideas are being conveyed with regard to AUB - its role in perpetuating an elite and its role in political development, the relationship being that elite perpetuation is a factor in political development. Although the relationships are not made explicit except in the qualifier of entry 4, they can be supplied fairly readily in entries 1 and 3 where the eye reads straight on in display and qualifier respectively. (In entry 3, elite perpetuation and political development might possibly be perceived as being in an 'and' relationship, and those reading from lead term to display, which conveys that the study deals with the influence of students' social class on AUB, might be initially puzzled.) But in entry 2, the break in the middle of the pair of interrelated ideas is felt to have a seriously disruptive effect.

Rotation effects in more complex subjects seemed another matter for investigation, although in the examples discussed during the training period the practice was generally to test selected rotations only.

3. Intellectual structure of descriptions. PRECIS indexing involves the twin processes of analysis and synthesis. Breaking the subject down into its constituent concepts (with the attendant problems outlined) forms part of the process of analysis. Another part of this process is the recognition of dependency or interdependency of the concepts one to another. Synthesis then involves setting down the concepts in a 'concept string' according to the rules of the system and coding by a set of role operators. The aim is to ensure that a given subject would always be represented by the same words in the same order.

Discussion started at the stage at which the subject is taken as given, but there is of course a prior stage of subject analysis leading to the initial statement of subject. Interpretation is inseparable from subject analysis and, whilst there are valid and invalid interpretations, there may well be more than one possible interpretation. Factors such as subject background probably have a substantial bearing upon interpretation. We became aware that the structure of the indexing language might also affect interpretation by predisposing the indexer, in terms of the questions the indexing language requires him to ask, to one interpretation rather than another, without due regard to validity.

The point is that determining the subject of a document is a process of imposing structure on its ideas; to describe or explain is, implicitly or explicitly, to structure. Thus a natural language subject statement has structure before the indexing process begins. The danger is that in breaking down such a statement, particularly if it is at all complex, and in reconstituting it in a manner prescribed by the system, the original intellectual structure may be altered. In this case both the overall meaning and the meaning of individual terms (as determined by context) would be changed, and the document misrepresented. We therefore wished to make a closer study of the intellectual aspects of the system.

A passive rather than an active 'sentence' ordering of terms is recommended in the PRECIS system, a linguistic 'transformation' from one to the other being considered non-distortional. It is a moot point whether the same actual message is conveyed, as contrasted with the same subject matter of message, and views as to the importance of this distinction will vary.

Beyond this, the main structural characteristics, embodied in the role operators, reflect systems thinking,

Operator	Meaning
(a)	Form - physical, e.g. Microfilm, or narrative, eg Journal
(b)	Target - e.g. For managers
/	Field membership, quasi-generic relationship
,	Difference - concept used adjectivally
(p)	Subsystem, structure, material
(q)	Property, percept
(0)	Study region, sample population
(1)	Viewpoint, perspective
(2)	Active system
(3)	Effect, action
(4)	Key system
(5)	Discipline
(6)	Environment
(v)	Coordinate concept
(w)	Coordinate correlated concept
(x)	Coordinate subject in same document

In general, the terms 'comprising' the subject are fitted in two ways into the concept string. 'Delimiting' terms occupy fixed positions at head and tail of the string (information concerning the problem as a whole), and the problem proper is expressed by means of 'flexible' (ie not tied to content of terms) coding in a central position in the string. Relationships between terms may be left to inference from relative position (eg difference, whole/part) or expressed verbally (eg in the string 'Universities. Manpower planning. Function of digital computers' - 'Function of' is an example of a device known as the 'gate').

Now to adopt a systems approach, as Runciman[4] points out, is to say little more than that there shall be formal structure. He suggests that saying 'look at it like a system' is rather like saying 'talk prose'. He goes on to say that appropriate use of this approach depends upon valid reduction in which one set of phenomena is accepted as identified (isomorphic) with another (in our case this would be contents of documents and contents of descriptions). He argues that components in the end must be related to meaning and thus structure to content.

By contrast, the PRECIS team advised us that in applying PRECIS, the syntactic structure of a subject was to be treated as quite distinct from its semantic content and that, in this way, results both reliable and consistent with document meaning would be achieved. Many of our subjects, however, were open to several interpretations and several representations in terms of the PRECIS role operators, when questions of meaning were left out of account in identifying 'systems' and relationship amongst them[5]. We were taught a number of unwritten supplementary rules associated with the role operators, but these were aimed more at ensuring helpful collocation than accurate representation of documents. It seemed to us that ultimately consideration of meaning of 'string' as compared with document was inevitable.

The problem with the role operators lies partly in the relative nature of categories such as 'system' and 'process'[6] partly also in the fact that the eventual configuration of a 'string' will vary according to how the starting point is defined in systems analysis, and a rationale for such decisions lies outside the logic of the PRECIS framework. In the absence of an explicit rationale, an implicit one will necessarily develop and become institutionalised. Such a rationale requires validation in relation to intellectual considerations. Inter-indexer consistency is not the only criterion.

A general question therefore concerned the nature of the implicit thinking underlying the implementation of the PRECIS role operators and the appropriateness of this thinking in relation to our subject field.

In summary, our first spell of work with PRECIS raised the following questions:

(a) How far could descriptions, not merely coextensive with documents but formulated in the author's terms and at the level of the document, be combined within PRECIS with grouping of documents in relation to broad subject headings;

(b) How far could 'fracturing' of complex ideas and procedures for manipulating the individual elements be combined with non-distortion of an author's ideas;

(c) How far could the length and complexity of descriptions formulated at the level of the document, the degree of fracturing, and the procedures for manipulation, be combined with non-ambiguity;

(d) Were particular assumptions about meaning implicit in the information handling procedures, and hence in the structure imposed upon the documents, or could the procedures be regarded as independent in this respect and if necessary be combined with alternative ways of defining the content of documents.

Optimising the system

The next phase of work was one of independent experimentation. The aim was to find ways of using the machinery of the PRECIS system to optimise index entries in relation to our objectives. Our exploration was entirely open ended, and the following section represents a roughly chronological account of it. Broadly speaking, we started with the problems of handling individual documents, went on to the problems associated with the generation of subject headings, and then looked at the relation between the two.

(1) Handling of individual documents

Our central concern was non-distortion. We would stress that a number of features of the PRECIS system take account most effectively of possible ways in which our material might be misleadingly represented, for instance the ability to distinguish between a 'convenience' sample and a population in which the researcher is centrally interested. Unfortunately, in some other aspects, the decisions built into the system are less appropriate. We considered first the most general aspect of the structure of descriptions - the PRECIS instruction to prefer passive to active order.

Independent experimentation showed that entries could be produced from an active order string just as readily as from a passive order string. It would seem therefore that consistency is the main consideration underlying this instruction. With the most simple problems this direction seems largely irrelevant (eg in the immigrant children problem the 'sentence' has no 'object'). The passive order instruction clearly fits problems of the pattern 'Influence of a on b'.

But, in more complex examples, the instruction may not necessarily be adequate. Thus one example we discussed was analysed and the following presentation was recommended by the PRECIS team:

4 Teaching style
q Effectiveness
3 Prediction
2 Contingency model
w related to
2 Authoritarian/democratic model
sub
3 Based on/Basis for
2 Leadership style)
) [view of] Group effectiveness as
v Situational) function of
variables

This presentation seems to imply that the purpose of the writer is to suggest how effectiveness of teaching style may best be predicted. An alternative focus would be that the validity of one model rather than another, as a basis for the prediction of group effectiveness, was supported by considering both in relation to a specific group situation such as the teaching situation. The alternative subject could also be expressed 'passively'. In this case, use of the passive order would produce a rough statement to the effect: (the view of group effectiveness taken by the writer and the relative merits of alternative models this view suggests) (supported by) (study of the prediction of effectiveness of teaching style). The difference of purpose and focus would be an important consideration by the user. Thus, 'passive order' as a guiding principle was felt to beg important questions when applied to some of our documents, as compared with the simpler examples used to teach the system.

Given that active order worked as readily as passive order, it seemed to us preferable to instruct the indexer to state first the theoretical framework from which the substantive content of a study takes its meaning, and to determine the links (active or passive) according to the structure of ideas in the document.

This, however, brought us into conflict with assumptions built into the role operators. In the first place, the role operators prescribe geographical location instead as the most general context within which the subject will be viewed. Secondly, a statement of theoretical framework will not necessarily be a simple label, but may require a statement as extended as that for substantive problem. We found that the range of conceptualisation in our documents stretched the role operators to their limits and sometimes beyond. They are geared to a more limited range of situations viewed in terms of 'influence' or 'role' models possibly, in association with geographical location as starting point, reflecting an expectation that subjects will be interpreted in relatively concrete terms. Additionally, the importance we attached to stating substantive problem in its relation to theoretical structure created a two-level structure in many of our subjects. This added a dimension for which the role operators do not allow. We needed, in effect, to distinguish between minor relationships (these within levels) and major relationships (those between levels). PRECIS is capable of handling considerable complexity, and does provide for the provision of information about treatment of a problem as well as for statement of problem. The situation in which a multi-level statement of problem is required is, however, one which the system does not recognise.

We felt that this might be the key to our problems of distortion and ambiguity. If this were the case, we should think of ways in which the role operators might be extended or reformulated to take account of this situation as well as those which they currently handle. On the other hand, it might be that if we could ensure that individual concepts and variables were handled without violence, the problems of overall structure of descriptions might be less serious. The handling of concepts and variables is interdependent with the more general question of the provision of subject headings.

(2) Generation of subject headings

The system requires an indexer to reduce the statement of a problem to its basic elements for the purposes of manipulation by computer. When a basic element is a multi-word term or phrase, any of the component words may be required as an access point, and special mechanisms are brought into play. We defined the 'basic elements' of our subjects as larger units than those assumed by the system. We hoped this might minimise the problems of overall structure of descriptions. This, however, had the effect of increasing the complexity of our units, with the effect that we still lacked a way of generating suitable access points without causing undue distortion or ambiguity in the process.

We did not directly address our problem of breadth of subject headings. We concentrated first on exploring the various PRECIS mechanisms associated with the provision of access points on the grounds that, with greater familiarity with this aspect of the system, we would be able to optimise the structure of our index entries. It would then be possible to tackle the problem of achieving some compromise between specificity and author's terminology in descriptions and broad controlled entry points without complicating factors entering in.

The PRECIS 'difference' device is designed to handle compound concepts such as 'socially disadvantaged preschool children' (in which successive adjectives give increasing specificity of a linear kind), which appear in entries as eg:

1. Children
   Disadvantaged preschool children
2. Preschool children
   Disadvantaged preschool children
3. Disadvantaged preschool children

It could not, however, generate the heading 'disadvantagement'.

This device also handles compound concepts with a different kind of linearity, the elements of which have, technically at least, a whole/part or possessive relationship, as in 'perceptions of role of language' which, if all terms were required as entry points, would be coded 'Language, role of, perceptions of' and would generate 'language', 'role of language' and 'perceptions of role of language' as lead terms.

The 'difference' device cannot always handle non-linear compound concepts such as 'Mothers' aspirations for children'. This combines two ideas ('aspirations of mothers' and 'aspirations for children', but such notions may well be regarded in combination as a single variable in a given study. The linear gate exists to handle this pivoting situation (eg role of/in), but the gate serves as a link between concepts and is not appropriate when it has the effect of 'fracturing' an idea unit, as in the following example:

Language. Role in socialisation. Preschool children
    Perception by mothers. Influence of social class ....

In this item it is the role of language in socialisation which is the object of the perceptions not language itself, and it is not the process of 'perception' but 'perceptions' which are considered. It seems virtually impossible to fracture and yet preserve all the links, and hence ambiguity may creep in when users supply missing links themselves, particularly if they read straight on from lead term to display. This particular area of study is fairly well represented in the literature and there would not be too great difficulty in supplying the missing links correctly, but it will be clear that the treatment of studies featuring unusual combinations of variables, might well give rise to ambiguity.

A further device which PRECIS offers is a KWOC-like facility which can be operated at the machine coding stage. This enables any term to be selected from a complex phrase and offered as a lead term. Thus language might be picked out from a given phrase and the entry would appear as:

Language
    Perceptions of role of language in socialisation ....

There is however the disadvantage that, associated with the principle that each term is established in the context of the next, there is the practice of presenting the phrase from which a term has been extracted in the display rather than in the qualifier. In one sense, this phrase is a more specific notion, but in another it represents the broader context within which the extracted term must be considered. However, leaving this aside, the effect is not inappropriate when the entries are read term by term as intended, but where the qualifier is lengthy this may not always happen, as in another example:

Individual needs. Indicated by comparison of response of
        white adolescents with response of negro
        adolescents to racial change strategies of civil
        rights organisations
    Perceived likelihood of satisfaction of individual needs as
        associated with orientation towards racial change

It is not individual needs alone which are 'indicated' by the responses to racial change strategies as might be supposed, but the relation between perceived likelihood of need satisfaction and orientation to racial change.

The PRECIS device known as the 'substitute', mentioned earlier, can also be introduced at the machine coding stage. A phrase may be prepared by the indexer summarising the content of a string up to any appropriate point and, once the next significant term has entered the lead position, this phrase will be automatically substituted for the discrete terms employed in previous rotations. Unfortunately, this device is available only in the qualifier, not in the display. Thus, in the role of language in socialisation example where a term at the head of a string is in lead position, it can do nothing to solve our problems although, with a later term such as 'social class' as the lead term, the qualifier could comprise a phrase such as 'Mothers' perceptions of role of language in socialisation of preschool children', and several sources of ambiguity would be removed. Another possibility is to prepare multiple 'strings' for a document, each of which maximises a different part of the subject.

The disadvantages with both substitute and 'multiple stringing' procedures are those of time and cost. It was questionable, in our view, that there would be any substantial benefit over manual indexing, since a considerable proportion of our material appeared to require handling in this way.

(3) Relation between handling of individual documents and generation of subject headings

On the basis of this work, we still hoped that with increasing familiarity with the strengths and weaknesses of the various mechanisms in relation to our material, and with considerable attention to their effect in different rotations, we would be able to produce entries with no more than a tolerable level of distortion and ambiguity. For an operational situation, use of the system would probably not be a viable proposition without some reconceptualisation of its intellectual basis in the light of some of the more recurrent problems, but we assumed that our findings would be fed in to a programme of system development work aimed at further exploitation of some of the features which make PRECIS distinctive from other systems.

Unfortunately there was a further aspect of the system with which it was more difficult to come to terms. We turned from the general problem of generating headings without disrupting the structure of ideas in our entries, to the problem of producing appropriate headings.

It is characteristic of the PRECIS system that alphabetisation and clustering is based not on the lead term but on the whole of the qualifier. Where subjects are treated in a manner coextensive with the document, this has the effect of bringing together in the same area of the index all material in which a given variable features as entry point. But, with specialised material at any rate, one tends to find that subjects are unique and that, particularly in later rotations, no qualifier is identical with any other. Experiment demonstrated that the effect is to produce as many sets of one as there are documents, and that these sets are in no helpful order, whereas broad clustering at least is essential, and an optional facility for creating subclusters would be highly desirable. For example, the following items form part of a sample all dealing in some way with social class*:

SOCIAL CLASS. Achievement in high school, students' intelligence, SES and occupational aspirations. Account taken of these variables in studying association of type of college with selection process and graduation rates
Methodology: sample of 1,253 college students in US. 045

SOCIAL CLASS. Association between political attitudes (economic/ non-economic liberalism/conservatism) and upward mobility of working class youth. Test of theory of anticipatory socialization
Methodology: sample of 138 male Catholic high school students in US. 026

SOCIAL CLASS. Association between SES and rate of infant development in period 8 months (Bayley test) to 4 years (Stanford Binet test)
Methodology: sample of 3,037 white children in US. 046

N.B. Capitals are used to indicate bold type.

SOCIAL CLASS. Attempt to reconcile conflicting theories of relationship between social class and success values (common and class differential values)
Distinction between wish for and expectation of success, and introduction of notion of perceived limitations with regard to self and opportunity structure. Methodology: sample of 465 white high school seniors in US. 043

SOCIAL CLASS. Comparison of relative influence of school (SES and program), peers and parents on educational aspirations of adolescents in US and Denmark
Implications for understanding of social mobility in terms of contrast between societies based on contest mobility and those on sponsored mobility. Methodology: sample of 2,327 high school students in US and 1,552 secondary school students in Denmark. 016

From this experimentation, it seemed that even if one were to optimise the handling of the individual descriptions in respect of the entry points required, one still would not achieve an index which, in its overall structure, would be convenient in use. There was one further aspect of the system we had not yet investigated which might have a bearing on this problem independent of the role operators and the procedures associated with them. This was the machine coding aspect, and a final spell of experimentation was devoted to exploring this.

Investigation of machine coding procedures

We had been concerned up to this time simply with preparing input in terms of role operators, and had not been involved with machine coding, which is a separate operation. It was not clear, therefore, whether the extensive PRECIS machinery could be employed to vary index format, or whether it was indissociable from the preferred format. Since the results of our work on individual descriptions represented at best an uneasy compromise, we wondered also if the machine codes could be used in conjunction with a different intellectual schema. One might argue that the problems resulting from the application of the role operators to our subjects as we defined them, even taking account of unwritten rules we were taught, were really such as to justify rethinking of the assumptions underlying the procedures for structuring descriptions. However, use of the PRECIS machinery to implement a rationale alternative to that of PRECIS would require that the machine coding procedures governing the generation of indexes were independent of assumptions implicit in the role operators.

It had been noted that the role operators did not directly govern the manipulations carried out by the computer and that instructions beyond those contained in the role operators were involved (eg to bring the KWOC device into play). It had also been pointed out to us that alternative machine codings could be associated with a given role coding to produce a different effect. There appeared to be two possibilities. One might either experiment with an alternative system of 'intellectual' coding prior to machine coding, or one might code directly for the machine, attaching definitions to the machine code. The latter was attractive in reducing time and effort, if it were possible, particularly since we had in mind the needs of a small scale operational situation in which the same person would be responsible for all stages of indexing.

We first consulted the PRECIS team as to the nature of the rationale upon which the system was based. Elements of several kinds of rationale had been introduced into discussion to legitimate different practices and, without clarification, it was impossible to tell what kind of modifications would be viable and what would be disruptive. We put this problem to the PRECIS team in the following terms:

"We are unclear as to the precise function, as you see it, of the role operators. In discussion these have been explained in several different ways: (a) purely as grammatical roles (and thus syntax is assumed to be a matter on which services using PRECIS (BNB or other) may decide their own preferences, eg active/passive, as to procedures to achieve consistency); (b) as a practical decision making guide (this might be termed a management device) to determine syntax, which we take to mean arrangement of words, in a string; (c) as a guide to syntax, associated with a conceptual model, which relates string structure to intellectual structure. It is clear that the role operators are an integral part of the system. The question is whether their use is mandatory in all circumstances. Our work tends to suggest to us that, in our circumstances, additional or alternative procedures from those explicit in the role operators are required, so as to structure input in such a way as to produce index entries which convey the ideas in our material as we understand it. The problems, as we see them, in applying the role operators vary depending on the definition one takes.

"As grammatical roles one may, despite the labels, interpret the operators in terms of subject, verb, direct object, complement (p.q) etc. But the syntactic structures with which these roles may be associated are limited in range. A subject may contain an indirect object. Where there is more than one 'verb' it may be necessary to bring out the distinction between inter- and intraphrase relationships, in such a way as to indicate 'main' and 'subordinate clauses' in a complex 'sentence', from the point of view of user understanding of the subject as a whole. The latter is a recurrent problem for us, and facilities for syntactic control at this level are in our view no less a part of syntactic control than 'concept/concept' relationships, and if one must choose are perhaps of prior importance.

"As a practical decision making guide, one would expect the rules associated with the operators to anticipate the questions which the documents would raise in the indexer's mind, and to prescribe the syntactic structure appropriate to the document in question. We have various additional questions we feel to be relevant. Also relabelling of the operators would help to avoid unintended interpretations of, say, (3). But even entering into the spirit of it, given a document in which there is more than one 'key system', the operators do not seem to us to indicate how the string should be ordered. Systems analysis may be called in to help. But even so, in translating the diagram into a string we seem to be asking 'is this what the author was getting at', nor 'does this fit the diagram'?

"Considering the role operators in association with a conceptual model, there is in principle an objection to 'a' model, in that to impose intellectual structure is to change the subject, in our view, unless the author's model happens to coincide with the particular model adopted. There is also the serious risk that by learning to think in terms of a given model, one may also be influenced in document analysis to interpret documents in this way rather than trying to view their contents as the author intended. As we understand it, systems theory (as distinct from systems analysis) has now given way to the notion of deep structure as the underlying conceptual basis of the role operators. Chomsky's ideas seem to have been challenged by a number of linguists, but this is not a debate we feel competent to enter into. But on more practical grounds, the labels attached to the operators are more or less the ones associated with the systems theory conception of them. Apart from a stated preference for a passive to an active 'transformation' where this is a relevant consideration, there seems to be no other explicit link to Chomsky's ideas. In terms of techniques in applying the role operators, it seems to us that the indexer must be looking for deep structures, not 'a' deep structure. Given that the structure of ideas in a document has been recognised intellectually and expressed in a statement of subject before indexing begins, it is not clear to us how the indexer should apply the notion of deep structure, so as to relate the varieties of intellectual structure recognised in documents to preferred principles of string structure, by means of the role operators."

In the event since, as this account of our problems shows, we had some doubts concerning both systems theory and structural linguistics as a basis for appropriate handling of our material, we decided for working purposes to make the assumption that it was a management device we were dealing with. In this case it would be possible to regard the machine codes as independent of the role operators, and to work directly with the machine codes, on the basis of an experimental rationale of our own. We could at the same time investigate the possibilities of varying presentation as well as intellectual organisation of the index.

The main elements of our rationale have already been mentioned. In terms of expected index use, we questioned that it is meaningful in the social sciences, at the present time at any rate, to assume search requirements which can be precisely stated, or search strategies whose logic can be predetermined, as is perhaps the case with enquiries to which intellectual activity in the natural and physical sciences gives rise.

Our model is rather one of the academic in interaction with the literature. The process of using a search tool may be conceptualised not as independent of and prior to utilisation of the literature, but as part of the utilisation of the literature. Hence a search tool should allow the user to simulate searching the actual literature in all its essential aspects (ie to respond to the same cues and so to arrive at the same relevance decisions as he would if working on the literature itself). Consequently, it becomes important to relate substantive content to theoretical framework in descriptions. For the same reason, it is desirable to retain the author's terminology in descriptions. This leads to the need for controlled headings by which to group documents in a way convenient for searching. These should be broad 'area' headings so as to avoid arbitrary definitions of specific concepts which have a variety of meanings.

In analysing our subjects it was considered that the variables or concepts as defined by the author were the smallest 'unit' that is meaningful. These units might comprise a single element (eg role) or several elements (eg mothers' perceptions of teacher role). People or setting might also be central to a statement of subject, characterising the population under investigation, but this would not necessarily be the case. The 'units' should be linked in terms of the relationships amongst them specified by the writer in defining a given problem. One might have a further more 'major' relationship between sets of variables as well as the relationships within a set of variables. Because of the non-linearity associated with the 'nesting' of minor and major relationships, we decided to confine manipulation of entries to major breaks in meaning.

We had a fairly clear idea of the kind of entries and the kind of index we wanted to generate. We devised guidelines for the preparation of descriptions in terms of major 'idea units'. These idea units could be coded by PRECIS mechine codes and their order of presentation changed as required by means of the PRECIS 'gate'. We intended to use the KWOC device to generate headings but were unable to combine use of the KWOC with control of the rotation of our idea units relative to access points. We wished to retain all our idea units in the display, thus not to impede clustering. The machine codes did not afford this facility. We could only achieve this effect by locating all access points at the head of the string, and access points which related to idea units occurring late in the string were then, in lengthy entries, quite widely separated from their context.

The reason lay in the structure of the machine codes. Certain positions in the machine codes, we knew, governed layout of elements and typography in the index entries. We assumed, therefore, that by appropriate coding we would be able to combine use of the KWOC device with control of the layout of entries we required. Unfortunately, we found that the relevant positions in the machine codes carry several instructions simultaneously. Thus, the designation of any element in a string as an access point automatically generates the 'two line' structure (and most of the typographical features) of the conventional PRECIS index entry, which for us precludes clustering. Only in a first rotation $^{(A}{}_{BCD)}$ does the entry not contain a qualifier. There appeared to be no means, open to the machine coder, of breaking the preset combinations of instructions.

In a report submitted to OSTI in July 1971, we presented our findings concerning the appropriateness of the PRECIS system relative to our requirements:

"(1) Taken as a given, descriptions coextensive with subjects of documents involving the use of author's terms in natural-language-order phrases preserved as entities throughout, coupled with use of a controlled list of broader terms to assign subject headings as entry points.

"(2) Clustering of our material, because of its specialised nature, can only be achieved in the absence of data in the qualifier.

"(3) In avoiding the mechanisms by which a qualifier is generated by the computer (using the KWOC device), we produce many entries in which the meaningfulness of the entry relative to the entry point is not immediately obvious (ie wherever the concept or variable to which the entry point refers occurs towards the end of the entry).

"(4) Even if this were tolerable, providing that the internal meaningfulness of individual entries were preserved, entries under a given heading may be numerous and do not appear to cluster naturally in any meaningful way.

"(5) It would be possible to combine a sub-heading with the heading as a single unit, each unit bringing together a separate group of entries (eg 'Social class and A', 'Social class and B' etc.) but there is a higher risk of material being overlooked than

> when subgroups of items are clustered under a single main heading. Helpful ordering might be achieved, though not distinctively typographically signalled as such, by composing the entries so that the terms selected for this purpose, possibly introduced specifically for the purpose, come to the fore. This is liable, apart from the time factor involved, not merely to result in linguistic clumsiness, but to create a possible barrier to comprehension and to risk distortion of a complex idea. Additionally it would often be necessary to use differential phraseology for different entry points.

"There has thus been failure to find a way within present PRECIS facilities of achieving meaningful clustering within the overall index as well as efficient control of presentation of the individual entry, ie in such a way as to ensure both the meaningfulness of the entry in relation to each of its entry points and the internal meaningfulness of an entry (irrespective of entry point). This is particularly clear when there is an accumulation of items with a given entry point in common. Optimisation of the overall index structure appears to reduce the meaningfulness of individual entries, and optimisation of individual entries appears severely to limit the possibility of producing an index helpful to the user in overall structure."

We were forced to the conclusion that the PRECIS system is not appropriate to the nature of our material and the way in which we may expect users to approach it. Ours is a specialised field, and it is important to represent accurately what is distinctive about different work in the field, as well as presenting it in a way convenient for retrieval. This means that document descriptions are relatively long and complex. The intellectual scheme in the PRECIS system, which guides interpretation and description of documents, tends to produce an initial distortion effect. Without a substantial amount of individual treatment of items, the manipulation procedures by which index entries are generated are sources of further distortion compounded by ambiguity. In principle, the manipulation procedures could perhaps be employed, in association with a more appropriate intellectual schema, to generate an index which did not have these disadvantages. In practice, both the PRECIS intellectual scheme and its manipulation procedures appear to be associated with a more general implicit rationale. The effect is that any modification to one disrupts the other. There appears to be no possibility of optimising index entries to requirements other than those assumed by the rationale. The rationale is reflected right through the system to the machine codes by which instructions are passed to the computer.

We recognise that our work has a bearing on the general applicability of the PRECIS system, although our concern has been simply with effective handling of our material. Our intellectual requirements are not likely to be untypical of specialised fields (particularly those where a proportion of the material reflects the forefront of research) in either the social sciences or the natural and physical sciences. At a more practical level, the facility to produce indexes in alternative formats would seem to be a generally desirable feature in an all-purpose system.

Such implications of our work are for others to judge. For our part, the number of points on which we felt doubt caused us unease, and the absence of clustering was the final factor which forced us to conclude that we must search in other directions for appropriate information processing procedures. Since that time our understanding of the nature of social science information has increased and, had the decision not been taken then upon practical grounds, intellectual considerations would later have forced us to the same decision.

VW/RAC
27.4.72.

References

1. Detailed comparison of the entries for most items is impossible for two reasons: (1) the sociology of education team included rather more detail as essential to bring out points likely to affect user perception of relevance; (2) there appeared to be differences in interpretation of subject.

2. Item indexed in a previous exercise by the PRECIS team.

3. This item was indexed by the PRECIS team in an earlier exercise. 'Recruitment of elite' was amended to 'perpetuation of elite' to correct a simple slip; recruitment is clearly summarised in 'political development', and it cannot have been intended to omit the central notion of elite perpetuation.

4. Runciman, W.G. What is structuralism? British Journal of Sociology, 1969, XX(3), 253-265.

5. For interest, readers might care to try to express the following in terms of PRECIS role operators:
   K-ations for L-ing of A Y-ity in N's of O-ation between S's B-ed on X Y-ity and on V Y-ity by C-ation of P B D-ence of E, F, G's on H-ations of I's in J and M.

6. Eg 'Sociology' was said to be by definition a process but, in our view, is equally validly viewed in structural terms. Again, the distinction between operators (3) and (w) is said to be decided in terms of whether the relationship is in the observer's mind or objectively real, but all relationships are in some sense in the observer's mind.

III 2. (ii) Experimentation with a PRECIS-compatible system, July - September, 1971

(Extract from 'Indexing requirements: progress report, July, 1971. Report submitted to OSTI in July 1971.)

The OSTI/Sociology of Education researchers have devised and undertaken preliminary experimentation with a 'new-style' system which may be viewed as a variant form of PRECIS. The essential features are:

(1) The initial statement (in author's terms as far as possible) is broken down into a small number of 'ideas' (as contrasted with the more minute breakdown into individual 'concepts' in conventional PRECIS).

(2) Headings and subheadings are to be assigned from an agreed list of broad terms and in accord with one of a series of agreed procedures (see proposals in report on Furzedown meeting of Working Party).

(3) The various elements are coded to provide data handling instructions to the computer.

Preliminary work on a manual basis suggests that this would greatly reduce indexing time, would result in higher inter-indexer consistency and would produce a more effective index from the points of view of intelligibility, readability, access and non-distortion.

The 'chunking' procedure described in (1) above may be utilised within the present PRECIS system (see section in previous paper), but the proposed alternative methods of clustering and formatting can not be achieved with the available facilities.

It is felt that a system of this kind, both in its technical aspects, and with regard to the intellectual basis for information handling we are beginning to develop, may well be appropriate to social science fields other than the sociology of education. The system has something in common with the abstracting process as well as with indexing and classification, yet it 'is' none of these. It is essentially a hybrid. In terms of use it may permit of retrieval and yet the techniques of use will perhaps come close to those usually applied to 'current awareness' tools or in 'browsing'. It may perhaps be viewed best as intending to provide simply an aid to searching the literature.

OSTI/SEA - 17.9.71.
OSTI/SEA Working Party

III A. (iii)

[Experimental system:]
Draft guidelines for 'abstractors' and indexers
[September, 1971]

Stage 1 Specification of content of subject

A a. What theory, conceptual model, framework or concept(s) are used as the theoretical basis for the study?

b. Are they viewed in any particular way (e.g. comparison)? State in relation to a.

c. Are important conceptual distinctions made?

B a. What are the major variables and what is the relationship amongst them? This may be a generalised relationship or a 'thesis' may be presented (e.g. x rather than y). State relation to section A.

b. What additional variables are included and what is their relationship to the main focus of the study (e.g. analysis by)?

c. Are important distinctions, operational definitions etc. specified, in relation to the problem under study?

d. Are subjects or context (social, educational, geographical) of interest in the study for their own sake? If so state. If not, sample information (under D) will be adequate.

C a. Has the study any important implications or bearing on issues not dealt with in the study itself? State at the level of for what or for whom. (If the implications are discussed in detail this should be regarded as part of the subject and included in A or B above.)

D a. Source of data. State (a) Details of sample, including nature and numbers of subjects; data on them which may be helpful to user (e.g. age, religious affiliation) in relation to the subject of the particular study, should be given in parentheses; also country in which study was carried out. (b) Type of study (e.g. Longitudinal study). If secondary use of own or other data is basis for study, state 'reanalysis of data', giving brief details of nature of subjects only in parentheses.

Stage 2

Structuring of subject statement

Overall structure

The general structure is that of the previous section:

A Theoretical framework. How related to B Content of empirical investigation.	Structure varied according to whether entry point relates to A or B e.g. A. Basis for B. or B. Based on A.
C Implications (for issues not discussed in article but on which article has a bearing). D Source of data.	C and D always follow subject 'proper' in this order.

(Some articles deal of course purely with empirical work, others are purely theoretical, and there may be no particular implications for other issues.)

The overall structure is indicated visually in final entries by full stops between elements (A – D); D additionally starts on a new line, and will possibly eventually be printed in a smaller typeface. At this stage of indexing, for convenience only, the elements are all written starting on a separate line.

Structure WITHIN individual elements (A –D)

A and B Theoretical framework and substantive content elements both have their own internal structure; the principles are similar in each case, relating to the nature of relationships amongst concepts. Structure is indicated visually by means of punctuation 'less' than a full stop: semi-colon, colon, comma and parentheses are used differentially in relation to strength of relationships amongst concepts (c.f. Fowler).

The semi-colon is used when separate problems, issues, etc., are treated at the same level in the same article e.g. Investigation of association between X and Y; relation to findings of other research.

The colon is used where there is a complex set of relationships (i.e. more than a single focus, a relation between two variables or a simple correlation amongst several). The major relationship is stated first, more minor relationship(s) following after the colon. The colon indicates that although the relationships are stated separately for clarity, they are closely linked and the eye should read on, e.g. Investigation of relation between X and Y: P, Q and R as intervening variables.

The comma is used in a similar way but in particular circumstances. It is used as in the previous example for variables in an 'and' relationship. Another circumstance is when the relationship of a given variable to others is not clear, e.g. Investigation of relation between X and Y, with consideration of Z.

A further circumstance is when variables are included for completeness' sake or for contextual purposes, but will not normally feature as subject headings[1], e.g. Investigation of relation between X and Y, taking account of) Z; or
analysed by)
Theory relating P to Q, distinguishing between R and S aspects (dimensions etc.).

Parentheses are used for information more closely attached to, but still in a 'subordinate' relation to, other information. Particular examples are definitional, specificational or explanatory information, which must be tied to individual terms or phrases to aid the user's understanding, but which must not distract attention from the main structure, e.g. Investigation of relation between X and Y: P factors (a, b and c), Q and R as intervening variables.

C and D There is nothing complicated about either of these elements. C is normally a simple phrase, e.g. Implications for X. D takes the form

Source of data - Sample information. Type of study.

See first section for more information.

Links between elements, and sequence of elements (A - D)

Once the precise relationships have been appropriately stated within any element, the element remains intact[2]. As complete units, however, the elements may be ordered differently as index entries are generated.

If (AB) constitutes the subject 'proper', C (its implications for work in other areas) applies to the whole subject (AB). Similarly D (Source of data) may be said to be relevant to the whole. For this reason C and D are consistently stated as third and fourth elements. Their relation to the subject 'proper' is made clear by the initial words 'Implications...' and 'Source of data -' respectively.

The relationship of theory to substantive content may vary somewhat: empirical work may be intended to test theory, established theory may be used to guide empirical work, and so on. The link between A and B is thus specified individually for each item. It is also specified in two forms: that appropriate when the work is viewed from an empirical standpoint, and that from the theoretical standpoint, e.g.

Element A (theory relating p to q)
Tested in/Test of
Element B (empirical work examining relation between x and y)

When an index heading relates to element B, element B is stated first and the appropriate link is used, e.g.

x
Empirical work examining relation between x and y. Test of theory relating p to q.

When the heading relates to element A, the entry is structured as follows:

q

Theory relating p to q. Tested in empirical investigation examining relation between x and y.

Stage 3

Assigning subject headings

Procedures with regard to subject headings are still experimental.

The practice at the present time is to identify *major variables and concepts* as potential index headings but to exclude minor ones from consideration as entry points (but see reference 1 page 93). Such variables or concepts may be social/psychological characteristics, aspects of social organization (social units or social relationships), social processes or analytical constructs (e.g. norms, values). General theoretical or conceptual perspectives are also relevant in this context.

General headings appropriately subdivided (rather than specific subject headings) seem preferable - if necessary omnibus headings (e.g. 'Prestige and quality') - so as to set the subject headings at a level of generality beyond definitional conflict. Cross references will guide users who enter the index at non-used terms to the appropriate heading (e.g. Quality, *see* Prestige and quality).

Terms *descriptive of the 'real world'* are under consideration. These fall into people, social/educational context and geographical location categories. Where these are of interest *per se* they should clearly be indexed as such. Where used as indicators for a given concept, concept rather than indicator should be indexed. Where they are 'simply' subjects or setting the case is less clear, particularly in view of the possible resulting bulk of the index.

In view of cross cultural etc. differences it would again seem that, to the extent that such entry points are provided, general headings and sub-headings are to be preferred to specific headings, since the user must inevitably also be directed to references under a considerable additional number of related headings. The alternative, entry at a given term plus single reference to broad section where all relevant material is collected and organised under sub-headings, seems likely to be more convenient for the user.

The basic principle should perhaps be to ensure that documents cluster in a predictable and helpful way. At both levels a separate 'contents list' of headings and sub-headings, alphabetically indexed in detail (as is *Psychological Abstracts*' practice), may be a convenient aid to access to relevant items in the array of entries. the *International Encyclopaedia of the Social Sciences* is perhaps a closer model. Principles for generating sub-headings are for further discussion.

Stage 4

Coding for computer (or clerk)

An example may perhaps best illustrate the procedure.

A 'string' is prepared as in the following:

038 Schiller, Bradley R. Stratified opportunities: the essence of the "Vicious Circle". AMERICAN JOURNAL OF SOCIOLOGY, 1970, 76(3), 426-442, Refs.(27).

1 Opportunity structure + Subheading 1 Social class + Subheading

2 Theory that distinctive opportunity structures confront different economic strata.

3 Tested by/Test of

1 Achievement + Subheading 1 Disadvantaged + Subheading 1 Home + Subheading

2 Investigation of achievement (educational and occupational) of sons of families receiving welfare compared with that of non-welfare sons from comparable home environment.

5 Source of data - reanalysis of data (national sample).

The simple set of codes (1 - 5) carries the following instructions:

1 = Subject heading (subheading will be indicated after + sign, if appropriate).
(2 = Element of subject.
(3 = Link, run on to:
(2 = Next element of subject
etc.
4 = Implications.
5 = Source of data.

There will be instructions also as to format (positioning and typeface) - yet to be finalised but provisionally as follows:

HEADING
Subheading (?optional)
Elements of subject stated here..........................
possibly running on to further line(s). Implications....
....
Source of data - follows subject......................
on one or more lines.

Description of process - Each '1' represents a separate entry point (subject heading and sub-heading) and these terms appear only as headings.

In the case of each '1', elements below the entry point are in turn stated (reading downwards) until the last '2' has been printed. The computer should then return to its entry point and print in turn all the remaining elements of the subject (reading upwards). '4' and '5' data always occupy the penultimate and final positions in an entry respectively. In the case of '3', only data to the left of a diagonal stroke is printed when reading 'down' and that to the right when reading 'up'. At some entry points the computer will find on searching that there is either no data above its entry point ('1's at top of string) or below it ('1's at end of string). Each item will have an identification mark (e.g. running number).

Stage 5 Generation of index entries

In the case of the example given above, the following index entries would be automatically generated (and alphabetised):

OPPORTUNITY STRUCTURE
(Subheading)

Theory that distinctive opportunity structures confront different economic strata. Tested by investigation of achievement (educational and occupational) of sons of families receiving welfare compared with that of non-welfare sons from comparable home environment.
Source of data - Reanalysis of data (national sample). 038

SOCIAL CLASS
(Subheading)

Theory that distinctive opportunity structures confront different economic strata. Tested by investigation of achievement (educational and occupational) of sons of families receiving welfare compared with that of non-welfare sons from comparable home environment.
Source of data - Reanalysis of data (national sample). 038

ACHIEVEMENT
(Subheading)

Investigation of achievement (educational and occupational) of sons of families receiving welfare compared with that of non-welfare sons from comparable home environment. Test of theory that distinctive opportunity structures confront different economic strata.
Source of data - Reanalysis of data (national sample). 038

DISADVANTAGED
(Subheading)

Investigation of achievement (educational and occupational) of sons of families receiving welfare compared with that of non-welfare sons from comparable home environment. Test of theory that distinctive opportunity structures confront different economic strata.
Source of data - Reanalysis of data (national sample). 038

HOME
(Subheading)

Investigation of achievement (educational and occupational) of sons of families receiving welfare compared with that of non-welfare sons from comparable home environment. Test of theory that distinctive opportunity structures confront different economic strata.
Source of data - Reanalysis of data (national sample). 038

References

1. A list of exceptions is being compiled; it includes e.g. 'subject specialism' as a minor variable (since material on this topic is fairly scarce), and 'local/cosmopolitan' as a 'distinction'.

2. There is in principle no reason why manipulation within elements should not feature in our techniques as well as manipulation between elements, but we are not convinced that the increase in effectiveness would justify the greater complexity and effort in coding and processing. However we propose to experiment further.

III Selected papers

3. Development of alternative rationale

Our experience in working with the PRECIS system had led us to question some of the assumptions about users and their information implicit both in the nature of the system and in procedures for implementing it, even though PRECIS represents some of the most advanced thinking, in its area, in information science. Since our work could proceed no further without appropriate conceptualisation to guide it, our next task was to think through the theoretical foundations of the information handling process. This thinking spans the whole of 1972. It was carried on concurrently with more practical tasks, but forms an earlier section in this report because it is necessary to an understanding of the development of our ideas concerning the practical tasks.

The first paper was prepared by the research team in 1972, and describes a preliminary attempt, in cooperation with our Working Party, to explore the issues. It was circulated for comment by subject experts and information scientists. Their response is reported in the next paper. The Working Party also completed a paper in 1973 based on the same thinking, but their aim was somewhat different. They sought to give a succinct account of the ideas, highlighting the issues which, from the subject expert's standpoint, are particularly crucial. It was planned to hold a conference in 1973 in which subject experts and information scientists could discuss the different viewpoints which emerged in this exercise, but unfortunately, none of the information scientists invited were free to come. However, we had taken the opportunity to outline some of these alternative viewpoints as we see them at the present time, and this outline has been extracted, from the papers circulated to those invited, to form the final item in this section.

III 3. (i) Methodological issues and information research
1972

CONTENTS

ABSTRACT

PART I GENERAL METHODOLOGICAL ISSUES

- The need to study methodology
- Implicit v explicit methodology
- Nature of methodology
- Information research as social research
- Methodology and social science information research
- Comparison between 'hard' and 'soft' sciences
- Methodology and epistemology

PART II METHODOLOGY: EXPLORATORY STUDY IN ONE SOCIAL SCIENCE FIELD

1. The construction of a conceptual model
   - Purpose of a conceptual model
   - Sources of ideas for model building:
     - (a) Theories of knowledge
     - (b) Structuralism
     - (c) Theories of change
     - (d) Theories of semantics
2. The construction of a system
   - Nature of guiding model
   - Functioning of a system
   - Overall framework of the system
   - The representation of individual documents in the system
   - Comparison with other types of information systems
   - Evaluation of system
   - Practical viability

Abstract

The first section of the paper contains discussion of methodological issues relevant to the information field. The importance of an appropriate conceptual model is stressed. Quality of research is dependent upon goodness of fit between model and the phenomenon under investigation. The central phenomenon with which information research is concerned is described as the generation of knowledge - a social process. Information research is considered as social science research. The problems of investigating the information situation in social science fields are discussed. The model most widely used in information research (the 'hardness' model) is seriously inappropriate as a representation of the knowledge generation process in the social sciences. It is argued that a more appropriate model is needed. The second section of the paper is in two parts, dealing respectively with model building and system construction. The need for a reconsideration of methodology was felt as a result of doubts raised in user study in the sociology of education. The alternative methodology outlined has been adopted in an OSTI-sponsored project concerned with the development of a scheme for the intellectual organisation of material relevant to the sociological analysis of education. This work is described so as to represent a case study which illustrates the thesis put forward in the earlier section.

Part I General methodological issues

The need to study methodology

In deciding upon appropriate methodology for research, it is natural to turn for ideas and guidance to other work in the field. This was our first step in a project to investigate information processing problems in one social science field (the sociology of education).[1] The aims of the research are to characterise user requirements, and to select or develop information processing techniques appropriate to them. 'Basic' research was expected to afford help with the former aim, and developmental and comparative research with the latter. Whilst the bulk of both kinds of information research focusses upon scientific and technological fields, there was no reason to suppose that the methodology would not transfer to other fields.

User studies early in our project were carried out along social survey lines. Questions as to user requirements were framed in the light of the techniques of information processing known to be available. We held seminars in which users were given examples to study. Our objective was to arrive at a decision as to the 'best buy' or the most appropriate combination of characteristics in a retrieval system.

It became evident from respondents' comments and later group discussion that the range of options failed to take account of issues of central concern in our subject field,[2] and that this raised serious doubts as to their usefulness. In particular, the bearing of differences in intellectual perspective[1] on users' views about the organisation of a body of material was stressed. The effect was to draw a sharp distinction between intellectual and technical problems in organisation of material. Comments also suggested that exploration of intellectual questions should be a prior concern, and that this might well make some technical questions unimportant if not irrelevant.

---

[1] Sponsored by the Office for Scientific and Technical Information. Our particular concern is with document processing to produce an information tool for manual use, the procedures of which should be if possible computerisable, but not necessarily computerised.

[2] Another social science project reports experiences similar to ours, that of respondents questioning the appropriateness of the questions asked: 'One of the problems, however, of sending questionnaires to social scientists is that many of them regard themselves, rightly or wrongly, as experts in survey techniques, and some clients obviously preferred giving their views on the value and phrasing of particular questions to actually answering them!' (Line, M.B. et al. Experimental information service in the social sciences 1969-1971. Final report. Bath, Bath University Library, 1972). Platt discusses the mixture of facts, theory and values involved in surveys intended to aid policy making. She warns against allowing respondents' theories special status other than as data, but excepts the type of case with which we are dealing, namely 'only ... where the sample was one of experts might the fact that they held them [ie theories] confer any claim to a special status ...'. (Platt, J. Survey data and social policy. British Journal of Sociology, 1972, XXIII(1), 77-91.)

Our methodology, an attempt to have users characterise their requirements in terms of available options, was in our view unsuccessful. Its weakness can now be seen to lie in a failure to consider, in principle, the appropriateness of available information processing techniques to the characteristics of our field. It was a mistake, in other words, to assume that the criteria guiding the development of such techniques were valid, in nature, range and priorities, in a context other than that which their designers had in mind.

We believe this to be so even in the face of the body of evaluation research. This is because evaluation validates only within its own terms of reference. The relationship between different areas of information research ('basic', developmental, evaluative) is circular. Investigations of user requirements tend to draw upon a range of 'established' criteria[2] in order to characterise requirements. These criteria constitute a framework which guides collection and analysis of data. Questions are framed in relation to these criteria. Information processing techniques are designed to give corresponding levels of performance along these criteria. Evaluation procedures are designed to rate or compare performance along these criteria. As Kuhn[3] points out, not all circularities are vicious, but they can be a source of real difficulty. The problem does not go unrecognised in information science.[1] We interpret our discovery (that issues of central concern were overlooked) as indicating that our field has characteristics other than those of the fields for which the 'established' criteria were developed. These characteristics may be different, or may be present but of minor importance in other fields.

---

[1] Defined by T. Shibutani as 'an ordered view of one's world ... a matrix through which one perceives his environment', in Reference groups as perspectives. American Journal of Sociology, 1955, 60(6), 562-569.

[2] Lancaster, F.W. Information retrieval systems: characteristics, testing, and evaluation. New York and London, John Wiley, 1968.

[3] Kuhn, T.S. The structure of scientific revolutions. (International Encyclopaedia of Unified Science, Vol.2, No.2) Chicago, University of Chicago Press, 2nd ed., 1970.

It is tempting to dismiss the intellectual problems raised by the social sciences as 'talk about conceptual systems and generalities'.[2] On this view, one could 'call a halt to this approach, however interesting it might appear, because it had little to contribute directly to the hardware of an information system', and therefore such conceptual issues are outside the present scope of information science. However, Foskett[3] argues that we must accept and learn to handle the categories that have a place in the literature.

It is equally tempting to explain away inconvenient responses, such as our enquiries produced, by attributing them to users' lack of sophistication faced with the technical expertise which has gone into constructing the tools we provide for them. It has even been suggested that social science users constitute a barrier on the further improvement of their information services.[1] This puts us in the position of seeming to assert that we know what is important or 'good for users', and users do not. If sophistication of techniques precludes users from understanding our principles, we impose our views upon our users. But social scientists have not accepted our views. In the social sciences, the existing services have attracted a massive non-use[2] or 'misuse'[3]. It is certainly possible to explain these 'facts' in terms of user non-sophistication. Equally, however, assuming that users want information, it is as plausible to argue that the services are inappropriate to users, as that the services are too sophisticated for users.[4] We have chosen the former explanation, on the grounds that beliefs about what is appropriate represent prior assumptions. We take this to indicate a need to develop alternative or additional criteria as a framework to guide research concerned with social science subject fields.

---

[1] For instance, Schultz et al. note that, in indexing evaluation, one set of index terms is commonly considered as a standard against which other sets are compared. The criterion set is 'established by some mode of indexing that is generally considered to produce a useful, if not optimal, product'. But 'the validity of this type of criterion rests on a basic assumption, usually implicit, that, because a mode of indexing has been, and continues to be, widely employed by information services (eg. assignment of terms by professional indexers), it must provide users with an "acceptable" level of retrieval performance. Although criteria validated by acceptability may be adequate for certain purposes, they have serious shortcomings for general use ...'. (Schultz, C. et al. Comparative indexing: terms supplied by biomedical authors and by document titles. American Documentation, 1965, 16(4), 299-312.

[2] Brittain, J.M. Information and its users: a review with special reference to the social sciences. Bath, Bath University Press in association with Oriel Press, 1970.

[3] Foskett, D.J. Classification and indexing in the social sciences. London, Butterworth, 1963.

The fact that conflicting explanations of 'facts' have equal face validity detracts in no way from the quality of the available technical skills. It does suggest a weakness in the methodology of the research which guides us in the ways we use these skills.

---

[1] Line, M.B. The information uses and needs of social scientists: an overview of INFROSS. Aslib Proceedings, 1971, 23(8), 412-434.

[2] Line, M.B. Op. cit.

[3] Our early investigations revealed that some of our respondents systematically used the practice of cover-to-cover searching. (Swift, D.F. et al. Investigation into Sociology of Education Abstracts: report on first stage of project. (Report submitted to Office for Scientific and Technical Information.) Oxford, 1970.)

[4] Cuadra believes that we should minimize the need for information intermediaries because 'there is something very personal about information needs' and hence there is risk of transmission loss. 'My own view is that one must ... move the user's skill gently but firmly forward'. (Cuadra, C.A. On-line systems: promise and pitfalls. Journal of the American Society for Information Science, 1971, 22(2), 107-114.) Taylor comments: 'On occasion ... an uneasy voice asks if the systems we are designing may be technically too advanced (but socially maladjusted) for present use'. He sees the user as having two choices: 'wanting what he gets', or getting what he wants by other means. (Taylor, R.S. The process of asking questions. American Documentation, 1962, 13(4), 391-396.)

It also points to a need to reexamine the basis of our assumptions. It is for this reason that the sociology of education project has attempted to break out of the vicious circle and to develop a methodology which will enable it to do so. This has involved working with subject experts independently to formulate criteria which are both acceptable to subject experts and viable in operational terms. The focus is upon intellectual organisation of material. The exploration of methodological issues which our work has involved forms the subject of this paper.

Implicit v explicit methodology

Paisley[1] commented in 1968 that 'we now know how studies *should* be conducted', but draws attention to 'conceptual poverty' in information research. He notes that sound methods (eg follow-up in the case of low response rate) can ensure valid data, but valid data do not ensure valid conclusions in the absence of adequate conceptualisation.

We would describe the methodological point at issue here somewhat differently. The variables which the researcher includes in his study represent a conceptual model of the situation he is studying. Relevant variables cannot always be investigated directly and operational indicators will be taken to represent them (eg performance on IQ tests may be taken as a measure of intelligence). The researcher always has a conceptual model. The researcher may describe his work at the operational level (eg in terms of IQ scores), but even his operational definitions represent an implicit conceptual model (eg of intelligence). In other words, tests aim to describe 'something'. The particular questions asked (of all the questions which might be asked), and how they are asked, reflect a particular view of that 'something'. The validity of the researcher's conclusions (these may be characterisations of user requirements, test results etc), depends upon acceptance of his *model*, by others, as an appropriate representation of the situation he is studying. The validity of conclusions is questionable if the conceptual model is not sufficiently clear or explicit, because appropriateness of model cannot readily be judged.

---

[1] Paisley, W.J. Information needs and uses, in Cuadra, C.A. ed. *Annual review of information science and technology. Vol. 3*. Chicago, Encyclopaedia Britannica, 1968.

It is also questionable whether it is meaningful to think of methods as 'sound' except in the context of appropriate conceptualisation. An approach studied in connection with our early user studies illustrates the point. It was at first assumed that a 'democratic' approach (ie summing individual responses and ascertaining the wishes of the majority) would be an objective and unexceptionable basis for decisions as to the most appropriate form of service. It became clear, however, that our users had wishes not only with regard to their personal use of information, but with regard to the contribution of an information service to the development of the field in which they work. We may need to make a conceptual distinction between a set of individuals as an aggregation and the same individuals as a collectivity. Soundness of method and valid interpretation of data will in this case depend upon a prior decision as to the level of analysis (individual, group or institutional) at which the researcher is to work.[1] There is a serious risk in selecting methods independently of a conceptual model simply because they are widely used or accepted within information science as sound. In a very important sense the methods selected determine the characterisation of the situation under study. In other words, our methods may distort the conceptual model.

Soundness of method is often equated with objective measurement. A concern for hard data may lead the researcher to select his variables on the basis of amenability to measurement. This may mean that he overlooks variables which can be studied only indirectly, but which may be more important in terms of appropriate representation of the situation. We judge this to be a particular trap when applying methods found useful in studying information problems in science fields to those in social science fields. For instance, our particular concern is with the intellectual organisation of material. Hierarchical and faceted schemes, in which documents are grouped according to explicit characteristics of the subject matter with which they deal, could not be made to handle material in a way acceptable to our users.[1] Current work in automatic classification has an obvious attraction as an alternative approach. Such work involves analysis of the structural characteristics of a body of information such as abstracts or citations. Structural characteristics can be reliably identified and quantified, and thus lend themselves readily to analysis.

---

[1] For instance Schultz et al. (Op. cit.) collected data from a group of subject experts, asking each to respond from their personal viewpoint, and then noted the problems of comparing data from individual authors (indexing terms) with that from the subject experts as a group, as they later wished to do.

The hardness of the results is somewhat illusory. The variety of techniques available for such analysis makes it possible to achieve a variety of 'results'. Karen Sparck Jones[2] has pointed to the need for rationale to guide such work. Specification of the kind of end product required is needed to guide a decision as to an appropriate technique of analysis. There is a danger that the researcher may rely upon analysis of his data to tell him what the relevant variables are. All statistical techniques require acceptance of implicit theory, namely that the important characteristics by which a body of data may be described can be quantified. If he specifies his variables by means of such techniques, the researcher must accept the assumption that quantifiable characteristics of a body of information (for instance the factors identified in factor analysis) are equivalent to its intellectual dimensions. This could lead him to define the aim of classification as to group material on the basis of those characteristics of its subject matter which can be readily quantified.

---

[1] Austin notes that Ranganathan 'showed that it is possible to break any conceivable subject down into its component parts, and then restructure these parts into a standard pattern by referring to a decision-making model which determines the order in which they should be written'. He states that: 'Modern classification theory has virtually been built up on this relatively simple idea.' Unfortunately, if differences in perspective are disregarded, use of such a model in relation to our material affords only consistency of a mechanical sort; it does not afford consistently appropriate handling. (Austin, D. PRECIS indexing. Information Scientist, 1971, 5(3), 95-114.)

[2] Sparck Jones, K. Some thoughts on classification for retrieval. Journal of Documentation, 1970, 26(2), 89-101.

We would contrast this with an approach which relies upon independent rationale to specify the intellectual dimensions which are relevant, and then selects a technique which will analyse material in terms of these dimensions. Such dimensions will not necessarily be directly amenable to quantitative analysis. This does not mean that we cannot measure them, nor that we should not. These dimensions must instead be tapped indirectly by means of carefully selected operational indicators. The latter approach to data collection ensures that relevant variables will not be overlooked because amenability to measurement is a defining characteristic of relevance. It also takes account of validity as well as reliability of data. The approach may be summed up by Homans'[1] dictum: 'Let us make the important quantitative and not the quantitative important.'

Users in science fields may perhaps, as Cleverdon[2] suggests, view their literature in terms of the subject matter they study, and this may represent a commonly shared perspective. Scientists are said to have accepted taxonomies and typologies of the objects they study. In this case, it is reasonable to suppose that in studying the structure of a 'historical' body of information, one can at the same time tap the particular dimensions of content or meaning in which one is interested. In the social sciences, however, a different model is called for. Differences in perspective are central to social science investigation. The relationship between structure and content of information is therefore more complex. The relationship between information and user may also differ from that in the sciences. Such aspects of intellectual organisation may not be directly amenable to measurement, but cannot be left out of account in any rationale for social science classification. In fact, intellectual dimensions may be confounded with quantitative dimensions if we overlook or regard the former as uninvestigable because they cannot be directly measured. Methodology appropriate to our field should at least leave open the possibility of studying the nature of the relation between them.[1]

---

[1] Homans, G.C. The human group. London, Routledge and Kegan Paul, 1951.

[2] Cleverdon, C.W. The effect of variations in relevance assessments in comparative experimental tests of index languages. (Cranfield Library Report, No.3.) Cranfield, Cranfield Institute of Technology, 1970.

To summarise this section, then, it is not only important that the researcher's conceptual model be made explicit so that its appropriateness, and hence the validity of conclusions, may be assessed by others. An explicit model is important as an aid for the researcher himself, particularly as a safeguard against misrepresenting, or failing to view objectively, the variables in the situation he is studying. This approach to research is more fully and formally described in the next section.

Nature of methodology

The term 'methodology' is often misused to mean simply methods and techniques of data collection and analysis. We define methodology as the 'logos' of enquiry. This view of methodology is relevant even in the case of the so-called factual survey. Whether the intention is to describe or explain, research is guided by a conception of the elements significant in the situation under study. This will guide decisions as to procedures of data collection and analysis. Conversely, also, different procedures imply particular theoretical assumptions.[1] The researcher who chooses a given procedure for reasons external to his research (eg because it is a 'standard' approach) must accept the assumptions implicit in this procedure. To the extent that conceptualisation implicitly or explicitly excludes major defining characteristics of the situation under study, the resulting data will constitute a distorted picture. Facts do not speak for themselves.

---

[1] It is interesting to find Green, a researcher investigating the structure of urban communities, making a similar distinction between (1) social structure in terms of social rank, familism or ethnicity, and (2) spatial, residential clustering. Green points out that 'the advantages of methodological rigour provided by a statistical processing strategy are outweighed by the difficulty of giving theoretical meaning to components, factors or clusters once extracted ... and cannot be used as a substitute for theory, in identifying significant dimensions in the first place'. (Green, B.S.R. Social area analysis and structural effects. Sociology, 1971, 5(1), 1-19.) Maron, in the information field, also deals with the point. It is probably true that we had machines before we knew what to do with them. 'But a more basic reason that solutions to our problems have eluded us thus far has to do with the fact that our subject is very difficult because some of its key aspects are basically epistemological, having to do with the activity of knowing.' (Maron, M.E. Mechanized documentation: the logic behind a probabilistic interpretation, in Stevens, M.E. et al., eds. Statistical association methods for mechanized documentation: symposium proceedings, Washington, 1964. Washington, National Bureau of Standards, 1965.)

When we come to use the conclusions of research as a basis for decision making, data and conclusions to which they give rise (or more accurately explanations imposed upon data) have validity only in so far as the initial conceptual model appropriately represents the 'real' situation. Quality of research is directly proportional to quality of methodology, and quality of practical decisions to quality of research.

Riley[2] describes the conceptual model as 'the researcher's image of the phenomena in the real world that he wants to study'. This is a generalised image, excluding detail considered to be non-essential to an understanding of the basic nature of the situation. It is not an idealised pattern to which it is considered the situation ought to conform. The researcher develops his model to fit the phenomena as he understands them, drawing upon insights, ideas, tentative theories etc. suggested by observation or previous research. In relating concepts together in a model he is forced to define his concepts and to test the logic of his thinking. The model guides further research in that major assumptions, as yet untested, are readily identified. It also guides formulation of research questions, specification of variables and interpretation of data. In the light of findings the model will be adjusted for better fit with the phenomena being investigated, thus increasing its explanatory and predictive power.

---

[1] The researcher who studies personal diaries for perceptions of reality as representative of a culture makes assumptions about the nature of culture by choosing to analyse it in this way, namely that cultures are such that they can only intuitively be understood and cannot be 'observed' in data collected from a random sample of respondents, however large the sample.

[2] Riley, M.W. Sociological research. 1. A case approach. New York etc, Harcourt, Brace and World, 1963.

The model is a practical research aid, the function of which is to provide an essential articulation point between theory and real life.[1]

Riley goes on to outline the bases upon which the researcher then makes a series of choices with regard to research design and 'empirical research methods that will translate his ideas into appropriate operations'. As Riley points out, there is a 'rapidly developing reservoir of available techniques' from which the researcher may draw so that 'the design used in each study is tailor-made'. In other words, methods are regarded neither as sound nor unsound in themselves. They are more or less appropriate to the problem as conceptualised. Adequacy of conceptualisation may be defined as goodness of fit with the phenomena concerned.

Information research as social science research

We define the phenomenon with which information research is centrally concerned as the generation of knowledge. An appropriate conceptualisation of this process must guide information research. It is essentially characterised in terms of people interacting with other people via documents. In other words it is a social process. This view perhaps requires clarification.

We cannot agree with a view of information processing as a matter of technology (Fairthorne[2] refers to the 'engineering development of symbol manipulating systems'). Nor are we in accord with Foskett[3] that the central problem of information research is the behaviour and properties of information. Both

---

[1] We recognise that this raises certain problems in the philosophy of knowledge about the 'reality' of external social life and the extent to which any particular model may at any point in time be assumed to be identical to it. Our position is that the model is the means by which the individual imposes consistency upon (ie explains) external reality.

[2] Fairthorne, R.A. Morphology of 'information flow'. Journal of the Association for Computing Machinery, 1967, 14(4), 710-719.

[3] Foskett, D.J. Progress in documentation: 'informatics'. Journal of Documentation, 1970, 26(4), 340-369.

Fairthorne's view and the conception of information as having 'behaviour' and 'properties' would suggest that information has an existence which is independent and external to individual human beings. Both dehumanise the knowledge generation process, and the latter reifies[1] 'information'. In fact one may even question that there is anything more than human beings who have knowledge, and externalise and communicate it in a variety of ways. Their cognitive behaviour is as it were 'frozen' in documents, and acquires meaning only as other human beings impose meaning on it. The notion that a structure of ideas or ordering of knowledge 'resides' in a corpus of documents and can be 'discovered' is also misleading.[2] Knowledge is created as human beings interact.

We are also unable to accept Lancaster's[3] argument that information systems only tell of the existence of documents and do not 'change the knowledge' of the user. Simply in becoming aware of the existence of one document, the way in which a user views another document may change. We regard use of an information service not as means to an end but as part of the knowledge generation process.

The view of knowledge generation as a social process is one which we believe to be appropriate across all subject fields, including the natural and physical sciences. Even scientific 'facts' are from time to time thrown back into the melting pot and reinterpreted to produce a different picture (or conceptual model) of the 'natural order' of things.[1] That such revolutions in thinking are comparatively rare, and that intervening periods are periods of stability, accompanied by systematic accumulation of so-called facts within a generally accepted structure of ideas, does not invalidate our point.

---

[1] Reification is here used in the sense of 'vivification', a step beyond objectivation.

[2] cf Foskett, D.J. Op. cit. In a similar way the view that search strategies can be 'based upon the content of documents in the system rather than on the accidental features of language that he [the user] or the author chose to use' is misleading. (Bell, C.J. Implicit information retrieval. Information Storage and Retrieval, 1968, 4(2), 139-160.)

[3] Lancaster, F.W. Op. cit. Lancaster's argument seems close to Fairthorne's belief that 'we are not concerned with the subjects of discourse as such, only how they are talked or written about, who writes and talks about them, how often they are asked for in what terms and by whom, and so on'. (Fairthorne, R.A. Op. cit.)

The degree of stability in the social sciences is much lower, because the rate at which models are reformulated and rejected in favour of others more appropriate is much greater. Additionally the number of models current at any point in time is much higher. Nevertheless the phenomenon is the same. Models are created, and found in time to be increasingly inappropriate. We may see this as poorness of fit with the 'real world' situation they are intended to represent or explain. Another possible explanation might view appropriateness as a function of the ideological needs of a society at any point in time.

For practical purposes, in the context of information research in natural and physical science fields, it may be viable to ignore the social nature of knowledge generation. Scientists are able to achieve a high degree of detachment from the phenomena they study because of the 'concrete' nature of these phenomena. A brick-upon-brick view of increments to knowledge is not too unsatisfactory. It may also be possible to ignore change in the structure of ideas because of the relative over-time stability of scientists' explanations of their phenomena.

Neither a reified view of knowledge nor one which sees the structure of ideas as relatively static provides an appropriate model for the social sciences. The reasons are these.

1. Not only is knowledge generation a social process, but the phenomena studied are social phenomena. It is therefore impossible for the researcher to remain wholly detached from his

---

[1] Kuhn, T.S. Op. cit.

subject matter,[1] either in his actual research or his reporting of it. Others who study his work are thus necessarily brought into contact with the individual researcher as well as with the phenomena studied by the researcher. A further factor is that the user is himself in every day contact with such phenomena even if he has not made a detailed study of them. As a result of his experience, he cannot be wholly detached in his study of the work of others.

2. Each interaction brings change. This means that no piece of research and no contact with a document or document representation can ever be exactly replicated.
3. The kind of interaction involved in knowledge generation involves the process of cognition. The user must either translate his ideas into those of the researcher's conceptual model or translate those of the researcher into his own. The researcher has a similar relationship with the people he studies.

Any mechanistic notion of knowledge generation involving a labelling of 'information', so that it may be retrieved by a means analogous to the use of a slot machine (however complex the machine, the 'channels' are fixed), is an inappropriate representation of the process in the social sciences.

It is also important to avoid conceptualising the knowledge generation process in a way which psychologises it into individual cognitive processes, into concept formation. The development and use of social science concepts is better characterised as a historical process of clustering and re-clustering of meanings. This effect is produced as epistemologies develop and fall into disrepute when they are confronted by alternative epistemologies within a discipline. A number of epistemologies may be emergent, current, or in decay at any point of time. Each will be associated with a particular grouping of knowledge.

---

[1] Friedman, N. The social nature of psychological research: the psychological experiment as a social interaction. New York, Basic Books, 1967.

Methodology and social science information research

We have taken the position that quality of methodology, both in basic and 'applied' research, is proportional to goodness of fit between the guiding conceptual model and the phenomena concerned.

We have argued that the central phenomenon with which information research is concerned in the context of social science subject fields, if not across all fields, is the generation of knowledge. The generation of knowledge has been described as an ongoing social process, involving cognitive interaction between individual researchers (or indeed thinkers and writers in general) and other individuals such as users (also thinkers).[1]

We consider first in more detail the nature of a model appropriate to social science information research, and then go on to discuss the implications for methods of research.

The interaction we have referred to is by no means uniform in character. Blum[2] notes that most discussions of sociological analysis mistakenly proceed as if a given term could be isolated from the tradition in which it functions. Individuals tend to talk from within a tradition or epistemology which they have assimilated and which they assume as a necessary condition of an adequate response. Communication depends upon whether others 'speak their language'. Blum uses this point as a base from which to criticise the hypothetico-deductive approach for which the natural sciences provide a precedent. The positivist argument is that such an approach is helpful in treating the world concretely, inducing order, looking for 'correct form', transforming the way in which the audience views the world; it is normative. Blum argues that this is antithetical to sociological analysis. He regards it as destructive of sociological insights to attempt to overcome the multiplicity of definitions that locate sociologists in their different perspectives. He is not interested in such a form of explanation at the expense of all that is distinctive in doing sociology.[1]

---

[1] We recognise the relevance of 'usability' factors, but view them as dependent on decisions concerning intellectual issues.

[2] Blum, A.F. Theorising, in Douglas, J.D. ed. Understanding everyday life: toward the reconstruction of sociological knowledge. London, Routledge and Kegan Paul, 1971.

We do not have here a simple consensus on the normative side of the fence versus infinite multiplicity on the other side. McHugh[2] notes that there are several epistemologies of positivism, for the reason that there is no single way of putting it into practice. The individual observer is limited by his own perceptions. When he describes, he does not simply 'cite the concrete practice'. He is engaged in the analytic observation or conception of the concrete according to some rule which enables him to distinguish between, say, mathematical notation and schizophrenic babble, which are concretely identical. There are various formulations of such rules and depending on the procedure adopted, so even the positivist's conceptions will vary and hence, also, the meaning of the terms he uses to describe the world.

---

[1] Wilson concentrates upon a distinction between normative and interpretive approaches which is closely related to that made by Blum. We have used this kind of distinction for our argument because it illustrates especially clearly the trap before the information scientist if he oversimplifies epistemological issues in a search for consensus. (Wilson, T.P. Normative and interpretive paradigms in sociology, in Douglas, J.D. ed. Op. cit.)
Some sociologists feel able to argue away perspectives competing with their own. Fallding, for example, widens the definition of functionalism to include all sociology. (Fallding, H. Only one sociology. British Journal of Sociology, 1972, XXIII(1), 93-101.) Most would agree that schisms exist such that perspectives focus upon different 'things' despite overlap in the words employed. Bernstein has recently analysed the basis upon which perspectives differ as follows:
(1) Those who place the emphasis upon the problem of order as against those who place the emphasis upon the problem of control;
(2) Those who place the emphasis upon interdependence and dependence, as against those who place the emphasis upon conflict and voluntarism;
(3) Those who place the emphasis upon how social reality is constructed out of negotiated encounters with others, and those who place the emphasis upon structural relationships;
(4) Those who emphasise the need to understand the everyday practices of members and the assumptions which make the daily practices work, and those who set up observers' categories and observers' procedures of measurement by means of which they reconstruct the constructions of members.
(Bernstein, B. Unit 17, in Swift, D.F. et al. School and society course. E.282. Bletchley, Open University, 1971. (Restricted circulation at time of writing.))

[2] McHugh, P. On the failure of positivism, in Douglas, J.D. ed. Op. cit.

We must accept therefore, as an appropriate representation of the information situation, that there are many different ways in which both our authors and users view the world, and that the same words may represent quite different conceptions. We must then accept that any attempt to translate statements of problems under investigation into a common 'language', for the purpose of information retrieval, is liable to be counterproductive. The information loss will often be almost total. Our model must therefore be a complex one allowing different brands of both positivism and antipositivism to express themselves without constraint. The nature of the knowledge generation process will vary in character depending upon the 'language spoken'. New 'languages' will come into being as new epistemologies emerge, and others will fall into disuse as disvalued epistemologies decay.

When we consider the methods of investigation we should employ, we face a general problem in the study of social behaviour that, even in a neutral role, the researcher influences that behaviour simply by his presence in the situation.[1] We have also noted the problems that arise at the best of times from assumptions built into a piece of research which condition its findings.[2] This effect is displayed as much within in vivo as with in vitro investigation such as automatic classification. As in any social research the researcher, however objective his methods, cannot achieve detachment from his subjects.[3] Where there are 'language' problems such as we have described, the aim of detachment will in fact detract from the quality of the research rather than enhance it. If the researcher is not to misunderstand the problem under investigation, he must be able to characterise it in its own terms.[1] To do so requires understanding of the different 'languages' used and the ability to identify with those who use them. Detachment would preclude this possibility. This means that a methodology involving interpersonal interaction, rather than a formal researcher/subject relation, is essential.

---

[1] Cf Roethlisberger, F.J. and Dickson, W.J. Management and the worker - an account of a research program conducted by the Western Electric Company, Hawthorne Works, Chicago. Cambridge, Mass., Harvard University Press, 1939.

[2] Polsby notes that 'What social scientists presume to be the case will in great measure influence the design and even the outcome of their research'. Such assumptions can even lead, as in his field of community power research, to researchers holding to their theory 'despite what appears to be contradictory evidence'. Examples are: 'The "false consciousness" argument, which holds that, when a social group violates an analyst's expectations, the group is acting "irrationally".' 'The "and-also" argument, which suggests that instances in which the analyst's expectations are not met are trivial or irrelevant.' (Polsby, N.W. Community power and political theory. New Haven and London, Yale University Press, 1963.)

[3] Friedman, N. Op. cit.

The need for interpersonal interaction with users does not however, in information research, arise merely from the requirement that the problems be defined in their own terms. The situation under investigation is itself one of interpersonal interaction, and this is interaction of a kind which cannot effectively be investigated except from a stance within the situation. We have to take account of differences in epistemology to understand cognitive interaction. A defining characteristic of an epistemology is that it is a way of thinking and does not itself exist in any entitative sense in peoples' minds or the language they use. It is taken for granted and its characteristics become evident only when inappropriate thinking brings it into question. In this situation the only approach to information research is to proceed by means of experimental processing of material, the procedures for which are successively modified as subject experts can then examine the results for distortion and pinpoint aspects of the procedures which introduce distortion. This enables the information expert gradually to clarify the principles and detail of the model which should guide practice, and to devise procedures (these may or may not involve use of the computer) which are appropriate to the situation. This means that a methodology of interaction involving cooperation, rather than the formal information scientist/user relation, is essential.

---

[1] MacIntyre discusses Winch's stress on this point and agrees that this is essential as a basis for social research. He points out that problems thus defined are in no way less amenable to objective study than those defined from a standpoint external to the actors as in the case of, say, Durkheim. (MacIntyre, A. The idea of a social science, in MacIntyre, A. Against the self-images of an age. London, Duckworth, 1971.)

Since defining what constitutes an acceptable solution can only be achieved by understanding the nature of the problems, an approach in which information researchers and subject experts pool their expertises, and jointly explore the problems and together create acceptable solutions, seems most generally appropriate in the social sciences. It has the advantage of dispelling any misunderstandings either group may have about the relevance of the expertise of the other. (The reactions of psychologists to efforts to 'improve' their information exchange practices appeared to stem from a misunderstanding of this kind.[1]) Our approach has the added advantage of a built in safeguard against unintended effects of change or innovation. Information processing based on research carried out by 'objective' techniques has no effective means of anticipating such effects, for the reasons we have given.[2] Yet such effects may be undesirable. Action based on social research (including information research) represents social intervention, and as such there is a responsibility for the appropriateness of such action. Our approach provides a basis for assessing the weight of different arguments concerning direction and extent of change when decisions are made and as they work out in practice.

We can now see that there are two senses in which the term 'objectivity' may be used. It may refer to the means (ie detachment on the part of the researcher so as to avoid biassing his conclusions and to ensure reliability), or to the conclusions (ie these should be valid or non-distorting of the situation under investigation). We would expect that in information situations where the 'languages' spoken are thoroughly understood, or are less numerous and various than in the social sciences, the trade-off between 'validity' and 'reliability' would be rather different. In other words it would be less probable that the objectivity

---

[1] Boffey, P.M. Psychology: apprehension over a new communications system. Science, 1970, 167, 1228-1230. The system referred to is outlined in Van Cott, H.P. National Information System for Psychology: a proposed solution for a pressing problem. American Psychologist, 1970, 25(5), i-xx.

[2] Line has advised caution on this score. (Line, M.B. Innovation resulting from research and development in the information field. 2. The user's view. Aslib Proceedings, 1970, 22(11), 559-569.)

(detachment) with which the researcher seeks solutions would seriously diminish the objectivity (non-distortion effect) of the solutions he produces. Where the risk of a distortion effect is minimal it would be foolish not to aim for the reliability that comes from detachment. Where the risk of distortion is high, the price is too great.

In common with researchers in all fields, information researchers may be expected to look to the 'classic' experiment as the ideal mode of research. It provides the criteria by which the absolute quality of a piece of research will be judged. But it is generally recognised to be an ideal which is unattainable, even in research on the physical world. In practice, we take it as a general principle that the researcher must exercise judgment in assessing the relative importance of alternative (and, as we have seen, possibly conflicting) criteria in the light of the characteristics of the situation he is investigating, and optimise accordingly. Where there is conflict amongst criteria, it is not only proper but, indeed, there is no choice but to optimise. There are no higher order criteria to which the researcher may refer to decide priority amongst 'ideal' criteria. The situation itself provides the considerations on which decisions as to optimisation must be based.[1] The appropriateness of the decisions cannot be justified in terms external to the research, but is relative to considerations internal to the situation. In other words, appropriateness of methodology, considered in terms of shortfall from the ideal, is determined by appropriateness to the situation under study.

---

[1] Our methodology has been to start from the academic concerns of users (how they work upon their subject matter, how they view their concepts and so on). This has suggested approaches to information processing, and successive experimentation has led to progressive formulation of the requirements to which information processing should conform. In this way requirements and processing techniques are gradually being brought into alignment by relation to the nature of the activity they are intended to support. The general procedure we have adopted may be seen to fall within that described by Blumer. 'Theory, inquiry and empirical fact are interwoven in a texture of operation with theory guiding inquiry, inquiry seeking and isolating facts, and facts affecting theory. The fruitfulness of their interplay is the means by which an empirical science develops.' (Blumer, H. What is wrong with social theory? American Sociological Review, 1954, 19(1), 3-10.) Subject experts are helping to decide upon the questions to be asked, as well as helping to provide answers. Polsby has discussed the pitfalls of panel work in some detail. (Polsby, N.W. Op. cit.) We are working with a small group of experts who collectively represent ideas in the field. The effectiveness of a panel for research into our problem, as we have defined it, does not depend in any important sense upon simple numerical aspects (proportional representation, size of group) of representativeness. We have been more concerned to ensure that within the group as a collectivity, or by additional means, all the important arguments will be raised and critically examined.

Comparison between 'hard' and 'soft' sciences

To clarify our argument further, we reformulate it as a comparison between the so-called 'hard' and 'soft' sciences.

The aligning of subject fields along a soft/hard continuum, now a commonplace, may be taken perhaps to provide a rough indication of the general way in which scientific information and scientist users are viewed.[1] The natural and physical sciences are characterised as hard because their concepts are considered to be relatively stable and unambiguous, their boundaries clearly demarcated, their research problems precisely defined, and so on. The social sciences present a complete contrast in all these respects. This is often attributed to their immaturity as subject fields.

The 'hard' characterisation of the natural and physical sciences is open to question as an appropriate representation of the 'real' situation. The problem of 'relevance judgments', for instance, remains unsolved.[2] Nevertheless, information science

---

[1] Cf Austin, D. Op. cit.

[2] It is now recognised that relevance is not a property of the document. (Kent, A. et al. Relevance predictability in information retrieval systems. Methods of Information in Medicine, 1967, 6(2), 45-51.) Alternative definitions suggest variously that relevance is a property of an information system, or consists in relatedness of documents, or of terms or topics within a document. Bar-Hillel is reported in 1964 as believing that we should be thinking in probabilistic terms. (Hillman, D.J. The notion of relevance (1). American Documentation, 1964, 15(1), 26-34.) Some of the range of personal and situational variables associated with differences in relevance judgments has since been explored. (Cf Cuadra, C.A. et al. Experimental studies of relevance judgments: final report. 3 vols. Santa Monica, Calif., System Development Corporation, 1967.) O'Connor notes that some people 'seem to believe that they are freeing the notion of relevance from obscurity and errors by explicating it in terms of satisfying information need'. (O'Connor, J. Some questions concerning information need. American Documentation, 1968, 19(2), 200-203.) Although Cleverdon (Op. cit.) discounts the problems seen by O'Connor (and others) in the con-

continued on page 120.

has built upon the 'hard' view of users and information. Requirements have been translated into standard operational terms and categories (indexes/abstracts, comprehensive/selective, retrieval/awareness/browsing, precoordinate/postcoordinate, hierarchical/faceted etc). Standard procedures are required of users. For instance, a particular method of retrieval (matching) has become established.[1] Indeed this has led some to define retrieval in

---

[1] Maron believes that 'the conventional search strategy ... is based on an invalid inference scheme'. 'The fallacy can be pointed out as follows: an indexer in the process of deciding whether or not to assign index tag I_j to document D considers the following sentence S:

> If document D satisfies the information need of a library user, then he will describe that need in terms of index tag I_j.

S is a conditional sentence of the form: "If X, then Y", where X = document D satisfies the information need, and Y = index tag I_j describes the user's information need. So we can schematize the transition from a user's request to the library response as follows:

If X, then Y
Y
Therefore, X

(The inference consists of two premises, one of which is sentence S, the truth of which is not now in question.)

To say that an inference is invalid is to say that it is possible for its premises to be true and conclusions be false. The above inference is clearly fallacious. We cannot even assert that the premises confer a degree of partial truth on the conclusion. It is not surprising that retrieval effectiveness suffers when based on an invalid search strategy.' (Maron, M.E. Op. cit.)

---

continued from page 119.

cept of relevance, Cuadra, writing more recently than O'Connor seems to be thinking along the same lines as O'Connor. He remarks that 'most systems are still badly designed, in terms of their capability for being used by persons other than the information intermediaries - the librarians and the information specialists. ... There is a limit to the extent to which one person - even a trained librarian or information specialist - can judge what a given user will regard as relevant to his information needs. (Cuadra, C.A. Op. cit.) Thus Taylor's suggestion still seems valid: 'messages ... stored in the system have no relevance to the inquirer's need until he initiates a question'. This suggests that we should study 'the "state of mind" necessary to decide that this item of information is relevant and that one is not'. (Taylor, R.S. Op. cit.) For the current state of thinking in this area, see Saracevic, T. Ten years of relevance experimentation: a summary and synthesis of conclusions, in North, J.B. ed. The information conscious society. (Proceedings of the American Society for Information Science, Vol.7.) Washington, D.C., ASIS, 1970.

terms of matching.[1] It has also meant that factors such as agreed definitions of concepts are viewed as necessary conditions for retrieval, because retrieval is thus defined.

In this way, the model with which 'hardness' is associated, with all this means, has come to be regarded not as a representation of the situation as it is in a given range of subject fields, but as a pattern to which all subject fields ought to conform.

This carries implications both for basic research and for information processing techniques. In any field both users and their activities are liable to be described in terms of conformity to the pattern. There is less concern to develop an alternative model objectively characterising the situation as found. Decisions as to information processing for retrieval, say, represent choice from the available range of techniques. Concern for appropriateness comes to take the form of imposing hardness, drawing boundaries and so on, even if arbitrary, to meet the assumptions implicit in the techniques. Users must be 'educated', or replaced by intermediaries, if they do not conform; the alternative is to bring services 'down' to the user.

Continuing with the example of retrieval, the problems inherent in achieving, even in science fields, the needed conformity to such requirements may partly account for the fact that natural language retrieval systems are by no means rare. Alternatively, it may indicate a changing view of requirements.[2] It is probably true to say, however, that even in such systems the expected mode of use is still matching. The computer has undoubtedly led to innovation. There has been some reconsideration of mode of use, eg cycling, iterative searching. But long standing assumptions die hard and it is possible that more radical re-thinking yet is appropriate. This would seem to be the implication, for instance, of Karen Sparck Jones'[1] plea for a rationale to guide experimentation in automatic classification.

---

[1] '... a document is "retrieved" when a match or partial match occurs between the formalized concise description of document and request. It can be seen that file searching is essentially a matching operation ...'. (Lancaster, F.W. Op. cit.)

[2] Cleverdon notes that 'users of a given system are likely to have a complex of needs'. He refers to 'areas of rigidity' in our systems such as the 'artificiality of a controlled index language', and suggests that we should 'build in the flexibility that will allow the system to cater to all types of users'. If we do not do so, 'we may produce systems that satisfy no one'. (Cleverdon, C.W. Design and evaluation of information systems, in Cuadra, C.A. ed. Annual review of information science and technology. Vol. 6. Chicago, Encyclopaedia Britannica, 1971.)

Even if we accept a crude soft/hard continuum as helpful to our thinking, concern with goodness of fit of the 'hardness' model would suggest that it is by no means wholly appropriate, for instance, to developing interdisciplinary fields in the physical sciences, and even some established fields in the natural sciences. When we come to most of the social sciences, the 'hardness' model is a gross misrepresentation of the situation. We fall below the threshold of reasonable approximation to actuality.

This should not lead us to regard the considerable number of social science users, and the particular nature of social science study, as 'sub-standard' or 'deviant'. We should develop and accept as valid an alternative model or models, redefine 'retrieval' in terms appropriate to the situation, and design information processing techniques and services in relation to our new conception of 'retrieval'. The important continuum, methodologically speaking, is not the soft/hard one, but badness/goodness of fit of model (be it formulated in soft, hard or some other terms) with the subject field we are serving. Judgment in his own terms by the user in the field is the ultimate test, not judgment in terms of assumptions external to the field, however valid in other contexts.

The main difference we note between the social sciences and other sciences lies in the assumptions about knowledge which are acceptable. The 'hardness' model is an oversimplification. Appropriateness of assumptions has a direct bearing on information research and on the appropriateness of decisions based on this research. A view of knowledge as an external 'concrete' reality is convenient and not too inappropriate in relation to the natural and physical sciences. It is seriously inappropriate to many social scientists.

---

[1] Sparck Jones, K. Op. cit.

We would in no way wish to suggest that information research and processing in social science subject fields cannot learn from the experience of information scientists in other contexts (this paper has focussed upon the specialised subject field, but a contrast might equally have been made with research and information processing in relation to general contexts, ie where users' interests range over a broad spectrum of subjects). In the simplest terms, we have to distinguish clearly between ends and means. Means found appropriate in one context are not likely to be appropriate in another if the ends are different. There are serious problems for the researcher concerned with social science fields in deciding whether experience in other contexts is applicable. Definitions of 'ends' will vary according to assumptions about information and about users. If these assumptions and definitions are taken for granted and not made explicit, appropriateness of 'ends', and hence of means (research procedures or information processing techniques), cannot be assessed.

It is possible to infer, as we have done, the nature of certain assumptions which appear to be widely held. (We have given reasons for rejecting these assumptions.) At the same time, research conclusions can be conflicting and accepted principles may be open to challenge (controlled v natural language) even within their own terms of reference. A process of rethinking appears to be in progress.[1] Even so, any alternative framework of assumptions or model which emerges will not necessarily hold for social science subject fields. If only to assess the applicability of such assumptions, and hence the validity of any reinterpretation of existing research evidence, those serving social scientists need to develop their own model. Such a model is in any case a practical necessity if they are to carry out their own research in the meantime with any confidence in the validity of their conclusions or the appropriateness of the information processing techniques they develop. The consequence of developing information systems guided by inappropriate assumptions is that such systems will inhibit the generation of knowledge to a far greater extent than they facilitate it. The cost of changing inappropriate systems must suggest caution, however strong the wish to take immediate practical steps and to initiate innovation or change.

---

[1] Foskett suggests that 'In our techniques for information control the time is ripe for the overthrow of the existing paradigm ...'. (Foskett, D.J. Classification for a general index language. (Library Association Research Publication, No.2.) London, Library Association, 1970.)

Methodology and epistemology

The definition of effective methodology we have offered is that it be appropriate to the situation under investigation. The researcher's model needs to be made explicit if its appropriateness, and hence the validity of findings, is to be assessed. Yet we have also noted that conceptions of a situation will necessarily vary. Thus, at one level validity is established if findings confirm the features of the model whilst, at another, the same findings may be open to question because the model itself is questioned. We have discussed so far the conditions bearing upon the first kind of validity. Addressing now the second kind of validity, we face epistemological issues similar to the order/control distinction noted by Bernstein[1] (whether social systems are the outcome of social interaction or whether they exist externally of individuals and impose restraint upon social interaction). Dawe[2] suggests that ultimately such distinctions represent doctrines and adherence to one or another is a matter of values.

In information research the contrast of 'doctrines' which is probably most often referred to is a user/system oriented one. There is a further distinction to be made amongst user-oriented stances, currently perhaps more generally favoured than system-oriented ones. These can vary widely depending upon the nature of the position accorded to the user in the scheme of things. In particular, stances vary according to whether information is viewed as the outcome of social interaction, or as external to users and requiring a system-relevant kind of behaviour of them. The externality view is the one represented by the 'hardness' model, together with its associated complex of beliefs about methods of research. The social interaction view is associated with the alternative approach we have outlined. Let us call the first an information orientation and the second a knowledge orientation.

---

[1] Bernstein, B. Op. cit.

[2] Dawe, A. The two sociologies. British Journal of Sociology, 1970, XXI(2), 207-218.

Neither orientation is independent of doctrinal belief. The information orientation, despite its association with an established model, does not afford an explanation of central problems such as relevance judgments. In this sense it can be said to represent adherence to doctrine. From this standpoint, the knowledge orientation can be said to be 'deviant', because it does not accord with the established model. Yet to the extent that the established model fails to hold, the knowledge orientation represents an alternative to, rather than a deviation from, the information orientation. It too, moreover, in so far as it can be challenged on grounds of inadequate representation of the situation, represents doctrine for those who adhere to it.

Taylor and Walton[1] have discussed the problematic nature of a fact/value dichotomy in the study of deviance. They use drug-taking as an example. The generally held value is that drug-taking is undesirable and the end is to stamp it out. The researcher's role is not, however, to participate in a means-ends relationship, and he will be able to demonstrate on factual grounds that the end is based on prejudice if he can show that certain forms of drugs are not harmful. The ends in that case do not take account of the facts. Cannabis-users may be considered deviant in the sense that society forces them into the position of a sub-culture because their values contrast with those of the dominant culture. But deviants are made not born. Deviance is a position in social structure, and what is deviant depends upon the way in which social structure is characterised. The point is that cannabis-users, say, cannot be termed 'deviant' on the grounds that they are irrational in their behaviour. Objective study simply shows that there are alternative rationalities which are not independent of values.

---

[1] Taylor, I. and Walton, P. Values in deviancy theory and society. British Journal of Sociology, 1970, XXI(4), 362-373.

The researcher must necessarily look at a problem from someone's point of view and hence is always open to criticisms of bias even if he makes his position explicit. He does not evade such criticism by adopting the position of the dominant culture and providing 'the opportunity of altering the odds in favour of their [the powerful] conception of normative reality being realised'. Taylor and Walton see as the problem 'not that ... we have to take sides, but that (given that, via method, we produce neutral "facts" and theories) we must not be so naive as to assume - indeed we must deny - that these explanations are neutral in their consequences for different "sides".'

This reminder has as much force for information research as for other social research. We show this by considering the way in which difference in orientation can guide the researcher to support different policies, with different consequences for users and system development. We then go on to discuss the problem of evaluating the consequences.

Line[1] believes that 'it is highly improbable that, for example, chemists organise information in totally different basic patterns from engineers', and hence can argue that 'there is a good case for rather more standardisation, on the grounds that a variety of arrangements of unknown utility is much more confusing to the user than one arrangement of unknown utility'. This is a similar belief to that reflected in proposals for the UNISIST network[2], which stress the desirability of 'rationalizing scientific information services into an integrated network'. The aim is 'to ensure that all that is known should be equally accessible to all who seek to know'.

---

[1] Line, M.B. On the design of information systems for human beings. Aslib Proceedings, 1970, 22(7), 320-335.

[2] Brookes, B.C. [Review of] UNISIST: study report on the feasibility of a world science information system, by UNESCO and the International Council of Scientific Unions. Paris, UNESCO, 1971. Journal of Documentation, 1971, 27(3), 216-220.

We may endorse this general aim and yet, if we reject the assumption of universal basic patterns, we may instead argue as Sharp[1] does, on the basis of 'the enormous difficulties in providing all-purpose, large-scale retrieval systems which will be all things to all men'. Equally practical considerations then suggest an alternative to the 'rationalisation' approach to realising our aim. This seems, to Sharp, to point to the probability that systems designed for limited fields, to be handled by the searchers themselves, is the right answer.

Thus a difference in orientation can lead to different views about policy, which can each be justified in relation to the same general aim.

In considering the consequences, Brookes[2] argues that a large-scale system such as UNISIST must give rise to disquiet. He questions the view of information 'as though its referent were some hard, cumulative, quantifiable, homogeneous, and transmissible entity as readily distributed and as undifferentiable as, say, electrical energy'. He notes references to 'control of science for the good of mankind' and 'control of the world's information services', which are made without indication of the nature of such control. He finds that 'the model adopted is primitive. In effect, the report implicitly assumes that what is good for chemistry and chemists is good for all disciplines and all mankind. The basic assumption is that any science ... can usefully be totally fragmented into bits of information which can be collected, sorted, and redistributed as required. For chemistry this model works reasonable well .... But few subjects have data bases of this kind ....' The report suggests that 'users need to be "liberated ... from scepticism about the information system's capacity to deliver". The implication is that reluctant users do not understand how to do their science.' This 'invites the retort: Has it not occurred to the authors of the report that your model of "information transfer" may not be generally valid?' In particular, 'Social scientists inspecting the UNISIST prototype model could reasonably ask what information scientists have been reading in the past twenty years'.

---

[1] Sharp, J.R. Where do we go from here? Aslib Proceedings, 1971, 23(1), 33-46.

[2] Brookes, B.C. Op. cit.

The doctrinal nature of the arguments on which assessments of consequences are inevitably based is clear from Brookes' comments. The 'conventional wisdom' assumes that the general 'good' is best served by adopting what is 'good' for chemistry as a basic pattern for all. But to argue that what is 'good' for chemistry is not 'good' for other disciplines is no less doctrinal. Even regarding the nature of 'the good' as problematic, our investigations must necessarily be inconclusive because the underlying epistemological problems remain.

Whilst however, in the final analysis, one must accept that there are alternative rationalities, it may be possible to establish that one set of assumptions leaves fewer 'loose ends' than others and hence, until superseded by one which leaves even fewer, is to be preferred. This can be a matter of investigation rather than assertion.[1]

In practice, however, any tool will fall short of perfection. It is therefore valid to ask whether the attempt is worth the effort. This means that we must demonstrate that existing tools are not only theoretically inappropriate but have practical disadvantages in use. These must be disadvantages such that (leaving aside that 'academics are being lazy or obtuse' in not using our systems or the explanation that 'we may be dealing here with some immutable characteristics of people'[1]) our efforts are non-productive or counter-productive in some sense.

---

[1] Possible reasons for theory choice include, for instance, accuracy, scope, simplicity, fruitfulness. Even these reasons represent values rather than rules of choice. Individuals who share such values may still make a different choice. One need not take the further step with many philosophers of science, and attempt to compare theories in terms of approximation to some ultimate notion of 'truth'. An example of the kind of justification we have in mind instead is that advanced by Kuhn in relation to his study of science: 'one set of reasons for taking the theory seriously is that scientists, whose methods have been developed and selected for their success, do in fact behave as the theory says they should. My descriptive generalisations are evidence for the theory precisely because they can also be derived from it, whereas on other views of the nature of science they constitute anomalous behavior.' (Kuhn, T.S. Op. cit.)

This cannot be empirically demonstrated (in the absence as yet of actual tools for comparative study and because users are technically unsophisticated) except in so far as we accept that all the evidence suggests that our formal systems do not 'perform' as well as informal communication. At the level of policy, however, we believe there are arguments which compel us to accept that, to the extent that we realise our policy, our efforts must represent a disservice to social scientist users.

Wagner[2] points out that 'Whether sociology moves towards theoretical unification is not decided by assertion; neither does it depend on statements of methodological "necessity". It is a matter of factual investigation.' We have to accept that 'Sociological theory is what theorists construct and consider as such'. There is considerable resistance to ad hoc standardisation of terminology. 'Premature insistence upon precision at all costs may serve to discourage the development of imaginative hypotheses, and without these sociology would indeed be the poorer.'[3] We should not take this to disvalue precision. Symbolic interactionists, for instance, are sometimes regarded as taking 'license for subjectivism', but even within this perspective there is recognition of the need for analytical concepts 'which are treated with care' as well as the more numerous 'sensitising' concepts.[4] The point is not to foreclose but to allow concepts to emerge from investigation of 'facts' as these are defined by any given perspective.[1]

---

[1] Line, M.B. Op. cit.

[2] Wagner, H.R. Types of sociological theory: toward a system of classification. American Sociological Review, 1963, 28(5), 735-742.

[3] Mortimore, G. and Enfield, R. What is sociology? 4. Their language and ours. Times Literary Supplement, 1968, April 4, 351-352.Cf also a complaint reported in a small scale study of patents examiners to the effect that 'mechanised retrieval forces the searcher into such a stylised way of thinking that he feels ... unable to explore and develop new approaches to solving his problem'. (Cornog, J.R. and Ellis, P.P. Patterns of thinking in searching patent applications by manual and machine-assisted methods. Journal of Chemical Documentation, 1965, 5(4), 215-225.)

[4] Lofland, J. Interactionist imagery and analytic interruptus, in Shibutani, T. ed. Human nature and collective behaviour: papers in honor of Herbert Blumer. Englewood Cliffs, New Jersey, Prentice-Hall, 1970.

By imposing 'hardness' we only succeed in creating a barrier between the user and his literature. The problem of a search will be determined by the user's perspective and his knowledge of the state of thinking within that perspective. We impede the user in locating relevant material if we attempt to 'improve' upon the present state of thinking. The system will not reflect his view of the field and will merely present a distorted picture of the literature.[2] The bias will not even be systematic and hence predictable, since the judgments involved must necessarily be subjective to the extent that information scientists do not engage in the 'factual investigation' needed to support them. Also, to the extent that the view of the literature we present may be accepted uncritically by those inexperienced in the field, further investigation would rest on unsound foundations. There seems therefore to be justification for an attempt to develop more appropriate tools on the grounds that effects such as we have described cannot be other than counter-productive.

---

[1] Blumer holds that 'The clarification of concepts is not achieved by introducing a new vocabulary of terms or substituting new terms - the task is not one of lexicography'. 'Serious attempts to grapple with this problem in our field' are 'alike in that the procedure is designed to yield through repeated performances a stable and definitive finding'. However even if such 'genuine definitive concepts of theoretic use can be formed out of the type of efforts I have been considering', Blumer suggests that 'thoughtful study shows conclusively that the concepts of our discipline are fundamentally sensitising instruments'. For this reason 'we should not assume too readily that our concepts are sensitising and not definitive merely because of immaturity and lack of scientific sophistication'. In developing such concepts 'We do not cleave aside what gives each instance its peculiar character and restrict ourselves to what it has in common with the other instances in the class covered by the concept. To the contrary, we seem forced to reach what is common by accepting and using what is distinctive to the given empirical instance.' Hence comes Blumer's stress on the necessary interplay of theory and 'fact' which we referred to earlier. Thus sensitising concepts, like definitive concepts, are brought 'more and more into line with what such study reveals'. (Blumer, H. Op. cit.) Cf also a discussion of sociological concepts in terms of 'constructs' and 'observables'. (Willer, D. and Webster, M. Theoretical concepts and observables. American Sociological Review, 1970, 35(4), 748-757.)

[2] A similar point is made by Tauber in relation to the chemical literature when he argues that information systems can most usefully reflect the state of knowledge about the subject matter under study. 'Imprecisions may have been introduced into information by oversight or by a very faithful rendition of the exact extent of knowledge'. He argues that information scientists should objectively reflect in their systems the exact extent of knowledge. (Tauber, S.J. Imprecision: problems for information processing. American Documentation, 1968, 19(4), 413-414.)

Given a need for more appropriate information tools, the choice of research approach is no less a matter of doctrine than is the nature of the model that (implicitly or explicitly) guides the work, whether it be basic research or development. In just the same way, it has to be accepted that there are alternative rationalities. Dependent on orientation, the researcher's assumptions will determine his view as to the approach which is appropriate to the situation he is investigating. We have shown, for instance, that acceptance of objectivity[1] as a criterion does not rule out the possibility that relatively incompatible research procedures can be justified in terms of objectivity. This is because alternative definitions reflect different values.

One major distinction in orientation in the information field with regard to this aspect of research may perhaps be described as 'practical' versus 'theoretical'. We take the 'practical' orientation to emphasise immediate applicability of research to the practical situation, and the 'theoretical' to emphasise a long term view of the value of understanding and insights into problems. The contrast comes out sharply in a comment by Fairthorne[2]: '"That's only theory!", they cry, from the back of the hall. "Only" theory, indeed! How do they think we reached the moon?' An intermediate position has been taken up by Paisley,[3] who has suggested that, in relation to the practical situation, there is a need for theories of the middle range such as were suggested by Merton[1] in contrast to abstract or general theories. Paisley contrasts theory with description, and notes that 'Purely descriptive study of any set of behaviors has its point of diminishing return'.

---

[1] Martyn stresses 'the need for objective measurement', at the same time noting some of the problems involved. (Martyn, J. Evaluation of information-handling systems. Aslib Proceedings, 1969, 21(8), 317-324.)

[2] Fairthorne, R.A. Innovation resulting from research and development in the information field. 1. A researcher's view: the detection of innovation. Aslib Proceedings, 1970, 22(11), 550-558.

[3] Paisley W.J. Op. cit.

The value conflicts underlying these different positions, which are in no way unique to information research, are outlined by Gamberg.[2] He shows that it is as fashionable in some circles to 'knock' concern for theory as it is in others to 'knock' pure empiricism. For the empiricist, 'facts have an intrinsic order that is distorted if the scientist approaches it with preconceptions of any kind ... "the facts speak for themselves"'. There is now 'realisation that the ordering and even the measuring of facts are dependent upon, and relative to, conceptual and operational ordering devices'. To the 'formalist', empiricism is the territory of 'vulgar bores who simply count heads'. Yet whilst 'On strictly methodological grounds, formal theory is unassailable ..., in terms of its applied relevance to any of the central problems of our time, it is of little significance'. Between these extremes there is a position in which 'theory appears to mean the notions that permit research to proceed', and 'while formulating propositions beforehand, conceives of theoretical constructs in operational terms'. This grounds theory in the type of data which have little relevance beyond the specific situation studied, and 'tends to lead to the same result as earlier empiricism'. For this reason, Gamberg believes we cannot find refuge in Merton's 'championing of theories of the middle range, because 'since we are never told what we are in the middle of (except somewhere between general theory and concrete research), the demand for this type of theory leads to an abnegation of explicit conceptual formulation',[1] and hence, once again, to findings which lack any general relevance.

---

[1] Merton, R.J. Social theory and social structure. Glencoe, Ill., Free Press, 1967.

[2] Gamberg, H. Science and scientism: the state of sociology. American Sociologist, 1969, 4(2), 111-116.

In our project, we are seeking to work within a theoretical framework which is applicable to a range of situations beyond the one with which we are specifically concerned yet, because it is applicable, is of practical assistance in guiding system development. This represents another kind of 'middle' position, avoiding the dangers both of head counting and of theory for theory's sake. Such an approach reflects an orientation which is no less doctrinal in its basis than others we have described. Justification again must lie in the number of loose ends we are able to account for by such an approach. The approach is essentially that stated by Fairthorne,[2] for whom practice and theory are 'aspects of the same reality'.

---

[1] Fallding deals in some detail with the 'refrain for "theories of the middle range"'. He argues that 'There is much to indicate that serious students of society are intent on explanatory theory as the end of their search. But if this leads to a rejection of conceptual or analytical theory it leads them into error of the gravest kind. ... the world of our experience has no semblance of order and, indeed, poses no questions unless we know the ways in which it may be differentiated. It is here that analytical concepts ... play their exploratory role'. He adds that 'there is a certain naiveté in imagining we will reach a general explanatory theory through special ones - unless we take care. Our special theories will have to be given in the same terms as one another if they are ever to be put together. ... It is the role of conceptual theory ... to help find those terms'. (Fallding, H. The sociological task. Englewood Cliffs, New Jersey, Prentice-Hall, 1968.)

[2] Fairthorne, R.A. Op. cit.

Part II Methodology: exploratory study in one social science field.*

1. The construction of a conceptual model

Purpose of a conceptual model

Our aim is to suit our methodology to the problem in hand. Our starting point is a definition of the information situation as a social one, and we focus upon the generation of knowledge. Consequently we must build a conceptual model which supports rather than contradicts these principles, and which provides helpful insights for the system designer. The task of the system designer in our view is to structure the information situation so as to enhance efficiency without seriously impairing effectiveness of knowledge generation (a 'trade-off' between the two aspects is assumed).

If we are to build an appropriate system, it must be on the basis of insights into the nature of the process to which the system is to serve as 'handmaiden'. We are aware of the differential meaning of knowledge. We have argued that artefacts of communication are no more than the external forms in which knowledge is momentarily 'frozen'. The burden of the argument in the first part of this paper is that the notion of the preexisting individual datum of information which is to be traced through the system is not viable as a basis for system design in our field. We are seeking insights into the nature of the process by which knowledge is created, because our system must participate in the process without disrupting it.

---

* We are deeply indebted to the group of sociologists and educationalists who have, over the past two years, met regularly to discuss the problems of intellectual organisation for literature relevant to the sociology of education. The project has benefited immensely by the ideas they have contributed to it. P. Gammage, University of Bristol; M.C.Grayshon, University of Nottingham; S. Hockey, University of Newcastle; R.K.Jones, The Open University; D.J.Oldman, University of Aberdeen; L. Watson (Chairman), Sheffield Polytechnic.

Sources of ideas for model building

(a) Theories of knowledge

Elias[1] distinguishes between philosophical and sociological theories of knowledge. In contrast to philosophical theories which focus upon scientific knowledge about nature, 'to insist on the social character of knowledge, is the great merit of a sociological study of knowledge'. Philosophers 'take their cue largely from the mathematic natural sciences in their classical form which they still tend to treat as an eternally existing and as a basically unchanging type of knowledge'. Knowledge is never completely autonomous, as this 'absolutist' view suggests, although Elias accepts that scientific knowledge has 'a relatively high degree of ... autonomy in relation to the structure and interests of different groups ...'. In general, knowledge may be seen as a 'nexus of symbolic representations, vested in language ... developed by men throughout the centuries ... which men gain in their societies', and not, as the philosophers view it, 'modelled on a single human being who is capable of orienting himself in his world entirely on his own'.

Sociological theories however, in Elias' view, fall into the opposite trap of 'relativism'. 'The latter take their cue largely from non-scientific types of knowledge about society which they ultimately treat as eternally changing like clouds in the sky in accordance with structureless changes of the groups where this type of knowledge is produced.'

In fact one simply has 'greater subject-centredness' or 'greater object-centredness'. Elias regards the concept of relative autonomy as helpful in characterising knowledge in different subject fields. He also questions the notion of successive phases of development in human knowledge, which do not recognise the 'long-term advances and regressions of human knowledge'. One needs a more flexible conceptual net.

---

[1] Elias, N. Sociology of knowledge: new perspectives. Part 2. Sociology, 1971, 5(3), 355-370. (For a discussion of the relevance of some ideas in the sociology of knowledge to the librarian see Holroyd, G. On the sociology of knowledge. Journal of Librarianship, 1972, 4(1), 48-56.)

In developing such a net, we may look to one sociology of knowledge, that of Berger and Luckmann.[1] This takes account, at least in a general way, of specialised knowledge in its relation to everyday knowledge, though it lacks the notions of relative autonomy and flux, to which we return after discussing Berger and Luckmann's theory.

Their theory is not concerned with universal structures, forms or perspectives on knowledge. It seeks rather to characterise and to explain how diverse structures, forms and perspectives arise, come to be accepted as valid, the reasons for their diversity, and the ways in which they are maintained or fail to be maintained in the face of competing structures, forms and perspectives.

Berger and Luckmann's starting point is the knowledge of everyday life. They describe a dialectic process between two types of knowledge: objective knowledge builds up as people interact with each other in routine ways and comes to be regarded as objectively real; subjective knowledge is acquired as people are then confronted by and apprehend this 'objective' reality.

The shared stock of knowledge which is 'taken for granted' as objective reality comes into being when habitualised actions become typified. This is accompanied by the development of 'roles' as people internalise such typifications and in turn 'realise' them. Essential to these processes is language, by means of which subjective experience is 'objectivated' and objective reality can become subjectively meaningful. The young child first 'takes over' the world as transmitted to him by 'significant others'. He does this by assimilating the categories by which they describe it; this provides an initial frame of reference which, as he participates in social interaction, is modified into a personal meaning system.

As knowledge of an everyday kind is transmitted from one generation to another, the situations which gave rise to certain kinds of behaviour no longer obtain. Explanations are then provided which legitimate this knowledge. This is knowledge about knowledge. Then too, as societies become complex (with the division of labour and an economic surplus), specialised groups come into being who develop their own differing stocks of knowledge, reality and symbol systems, and legitimations. The existence of these groups is supported by a secondary socialisation process which follows that of childhood. In this process the reality system associated with, say, a profession, is 'taken over' from the relevant 'significant others' and integrated with that internalised in childhood. Increasingly knowledge becomes socially distributed.

---

[1] Berger, P.L. and Luckmann, T. The social construction of reality: a treatise in the sociology of knowledge. Harmondsworth, Penguin Books, 1967. (First published in U.S.A. in 1966.)

With the segmentation of knowledge and the development of a variety of legitimations, there has to be some sort of integration at a theoretical level for people to make sense of it all. We are now concerned with realities other than those of everyday experience. An explanation of institutional order is said to constitute a symbolic universe. This too is socially constructed. As long as the situation which it explains persists, it is regarded as objectively real (indeed it may even be reified in a way comparable to 'natural order'). But, social situations are fluid, and theories tend to have progressively less plausibility as the situations they are intended to explain progressively change. Their status as theory then becomes problematic, and they will face the challenge of competing or counter realities.

It will be clear that, in the case of the specialist subject field, we are concerned with something lower down the scale than the symbolic universe. We have probably to deal with what Berger and Luckmann call cognitive sub-universes. To a certain extent these are worlds one may enter and return from to the everyday world, and they are clearly seen to be grounded in the latter. This apart, however, they may be considered to function as microcosms in a way comparable to the larger reality systems represented by symbolic universes. They have their own meaning systems, theoretical explanations and tests of reality. Members are regarded as significant others, each to the other, in maintaining the reality system as legitimate in the face of competing systems. There will be varying degrees of tolerance amongst competing reality systems.

Thinking at this level we avoid the problem of allowing for individuals' personal and possibly idiosyncratic meaning systems, yet at the same time focus upon the differential nature of meaning, which is of central concern in social science information processing. The range of individual differences can be seen to be structured according to membership of cognitive sub-universes. The differences which occur at the level of cognitive sub-universes are the basis on which we may work.

This theory steers us neatly between absolutism (knowledge as discrete from the individual) and relativism (the individual with all knowledge contained in his personal and idiosyncratic meaning system). Yet there are the problems noted by Elias, that we have no basis for detailed characterisation of knowledge at a given level either at a particular point in time (structural aspects) or over time (processual aspects). There are several lines of thinking we may use in building such detail into our model.

(b) Structuralism

Drawing upon Berger and Luckmann, but in line with structuralist thinking, Bernstein[1], for instance, has considered the school

---

[1] Bernstein, B. On the classification and framing of educational knowledge, in Young, M.F.D. Knowledge and control: new directions for the sociology of education. London, Collier-Macmillan, 1971. Elsewhere Bernstein notes that: 'From different sources, Marxist, Phenomenological, Symbolic-Interactionist and Ethnomethodological viewpoints began to assert themselves. Although there are major differences between these approaches, they share certain common features.

1. A view of man as a creator of meanings.
2. An opposition to macro-functional sociology.
3. A focus upon the assumptions underlying social order, together with the treatment of social categories as themselves problematic.
4. A distrust of forms of quantification and the use of objective categories.
5. A focus upon the transmission and acquisition of interpretative procedures.'

Contrasting the previous structural-functionalist approach with these approaches, he points out that the first emphasised the macro-structural relationships and tended to assume a normative system. '... the second approach takes as problematic the normative system and its acquisition, but it, itself, presupposes a complex structural arrangement which provides, at least

continued...

curriculum not in terms of 'traditional' subjects but as an element in 'educational knowledge codes' (which refer to the underlying principles which shape curriculum, pedagogy and evaluation), and has traced the implications for teacher/teacher and teacher/pupil interaction, thus clarifying the overall nature of the problems which confront policy makers.

Such work is based on the assumption that considerations of the structure of knowledge are necessarily intertwined with understanding of the processes by which knowledge is 'realised'. The notion of 'educational knowledge code' is the means by which Bernstein relates differential communication processes in the pedagogical situation with differential structuring of knowledge. Central concepts are classification (defined as strength of boundaries between areas of knowledge) and framing (nature of control over the generation of knowledge). Dependent on the relation between these elements, so patterns of control in social relationships will vary.

The central point of interest is that the ideas of structure and communication represent two ways of viewing the same phenomenon. It is by imposing structure that communication becomes

---

continued

initially, and often finally, the terms of locally situated activities.' This means that there is 'the issue of how we relate macro and micro levels of explanation ... the question of how theories which are based on very different assumptions are to be related'. Phenomenology 'does not show us how, it simply indicates a direction'. Such issues seem very far from the everyday world, but 'everyday activities carry within themselves the processes and practices crucial for the understanding of more general questions'. Thus, 'The levels, if they are to be usefully linked, must be linked at the substantive level by an explanation whose conceptual structure directs empirical exploration of the relationship between levels.' Bernstein's 'structuralist' concepts of classification and framing are concepts of this kind (NB the distinction between structural and structuralist is an important one). The kind of problem to which he applies them ('if we are to consider the relationships between schooling and society a crucial question becomes that of accounting for the constraints which limit the style educational knowledge takes for groups of pupils and students') has an analogue with the problem of the relationships of the concerns of specialist fields of research both with each other and with society at large. (Bernstein, B. Unit 17, in Swift, D.F. et al. School and society course. E.282. Bletchley, Open University, 1971. (Restricted circulation at time of writing.))

possible, since it is by relating information to other information that information becomes meaningful. If the actors in a situation do not have a common base of shared understandings, there is a serious risk that they may exchange utterances without effectively communicating. (The problems of the working class child vis à vis the middle class teacher have been the focus of much of Bernstein's work.)

It is not clear to what extent Bernstein accepts structuralist assumptions as described, for instance, by Piaget,[1] but the closeness of the analogue with the situation we are studying is justification for looking more closely at structuralism. Piaget holds that 'the notion of structure is comprised of three main ideas: the idea of wholeness [with the properties of conservation and stability of boundaries], the idea of transformation [which is subject to 'laws'], and the idea of self regulation [associated with the principles of reversibility, permanence and independence of end result from route taken]. These ideas appear to give a basis for a non-relativistic approach to the study of knowledge, which nevertheless concentrates on knowledge generation. For Piaget, a structure is a system of transformations. The laws governing transformations are not immutable. Rather, structures in being constructed give rise to the necessity of given outcomes, but in no sense 'contain' what is derived from them. Man, however, is more than a stage upon which events are enacted. By his 'reflective abstraction' he performs 'the operations which constitute the elements of the structures he employs in his ongoing intellectual activity'. Piaget argues that all social sciences, including sociology, yield structuralist theories. But he believes that 'the unconscious activity of the mind consists in imposing forms upon content' and that 'these forms are fundamentally the same for all minds'. This means that it is difficult in discussing the anthropological structuralism of Lévi-Strauss, to explain the 'advanced' logic of kinship systems of 'primitive societies' in relation to 'western logic'. He argues that these 'are finished systems .... What we want to know about are individual inventions'. The applicability of structuralist theories to our work, in the light of our particular view of knowledge generation as a social process, must remain an open question on the basis of Piaget's exposition.[1] His statement of the 'constructivist hypothesis' about 'the nature that underlies physical reality' seems to represent an epistemological position. However, as Glucksmann[2] points out, 'Despite basic similarities of orientation, there are important theoretical differences between structuralists'.

---

[1] Piaget, [illegible]. Structuralism, translated by C. Maschler. London, Routledge and Kegan Paul, 1971.

Runciman[3] disputes that there is anything distinctive about structuralism in either doctrine or method that is not assumed in all scientific enquiry. He suggests that, as in the case of systems theory, hope for a general theory creates enthusiasm.[4] However, to say 'look at it like a system' is rather like saying

---

[1] On this point Gardner believes that 'Piaget and Lévi-Strauss have invoked structuralism for different sorts of inquiries. Piaget has focused on the developing structures ... while Lévi-Strauss has elected to describe the different forms the structures can ultimately assume'. Again, 'Lévi-Strauss typically examines the relationships between individuals, while Piaget examines the action of one person on the world of objects'. He suggests that in both cases, 'while neither neglects subject-subject relations wholly, subjects are viewed primarily as human objects'. (Gardner, H. Piaget and Lévi-Strauss: the quest for mind. Social Research, 1970, 37(3), 348-365.)

[2] Glucksmann, M. Structuralism: a review article. British Journal of Sociology, 1971, XXII(2), 209-213.

[3] Runciman, W.G. What is structuralism? British Journal of Sociology, 1969, XX(3), 253-265.

[4] Attneave makes a similar point about cybernetic approaches and information theory, to which people 'reacted with an excess of enthusiasm', but which do not provide 'a ready-made solution' to all problems. (Attneave, F. Applications of information theory to psychology: a summary of basic concepts, methods, and results. New York etc, Holt, Rinehart and Winston, 1959.)

'talk prose'. It may be helpful as a method[1] in the sense that breaking down of the object under study and reconstitution in terms of essentially relational properties is model building, but there is nothing distinctive in recognising that a model is a construction. The point is to achieve 'valid reduction', and validity will depend on content,[2] not on any connection with the notion of structure except in so far as all sociological theories are structuralist. In view of the present stage of thinking in this area, we believe that Berger and Luckmann's conception of the knowledge generation process is more in keeping with our ideas, but feel able to borrow from structuralism to build in some of the detail we require without entailment to theory associated with it.

Berger and Luckmann's explanations of the social distribution of knowledge concentrates upon the relation between specialised knowledge, such as that of cognitive subuniverses, and everyday knowledge. In the context of social science information handling this is a relationship we must not ignore. We are interested also in differences between cognitive subuniverses and their associated reality systems. The search for underlying universals is a task for subject experts. We simply want to characterise reality systems as they are for the purpose of building our system. A first step is to make an analytical distinction between operations and structure, remembering that it is only an analytical one.

For our purpose, it will be helpful to consider the 'operations' aspects of knowledge in terms of the degree of detachment of users from knowledge in different reality systems. Amongst other things, the extent to which users of a reality system at any point in time are able to accept its knowledge as objectively real will determine how they operate in knowledge generation. The less detached they are, the more effective interaction will depend on a sharing of understandings with other actors.

---

[1] This distinction is similar to that between (1) theorising about biological systems or social systems and (2) the application of systems analysis to the design of complex engineering systems and associated work in operations research and cost benefit analysis. (Emery, F.E. Systems thinking: selected readings. (Penguin Modern Management Readings.) Harmondsworth, Penguin Books, 1969.)

[2] Cf the problem inherent for us in modern classification theory as characterised by Derek Austin, which was touched on earlier.

With regard to the structural aspect of knowledge, we may expect that the boundaries between reality systems at any point in time will be more or less permeable. This means greater or lesser possibility of meaningful integration of one reality system with another. The analogue here is with the notion of boundary maintenance. However, this is not a simple notion, since also to be considered is the nature of the structure of knowledge within reality systems. The 'outer' boundaries may be strong or impermeable, yet internal structure may be close knit, and the elements highly interdependent or permeable. Conversely, the outer boundaries may be weak or permeable yet internal structure may be loose knit and the elements largely independent of one another or impermeable, and hence 'objectively' or formally rather than subjectively and variably related. The more permeability there is within or across systems, the more the meanings of individual elements will be interdependent.

We regard as a matter for empirical investigation the 'rules of transformation' which may underlie the knowledge generation process. We hypothesise that the greater the degree of shared understanding required for effective interaction, the more permeable the categories employed within a reality system will be and at the same time the stronger the boundaries between that system and others. Conversely the more 'objective' the interaction, the less permeable the categories employed will be, and the less important to members of a reality system the maintaining of boundaries. However, the greater the non-permeability of categories, the less acceptable they will be to reality systems depending upon shared understandings, and the stronger the boundaries against them.

(c) Theories of change

Since we view knowledge generation as a social process, we may expect change over time as well as differentiation at a given point in time. We must therefore build a 'historical' dimension into our model. We believe that our system must mirror change, at least to the extent that change modifies the nature of shared understandings and thus affects use of reality systems, if it is not to disrupt the knowledge generation process.

We have noted the inappropriateness of the notion of brick-upon-brick accumulation of knowledge as a general conception of the knowledge generation process over time. This does not mean that the process is not patterned. Kuhn[1] presents an alternative characterisation of it, central to which is the notion of paradigm. A paradigm is 'the entire constellation of beliefs, values, techniques, and so on shared by the members of a given [scientific] community'. As such, it clearly has much in common with Berger and Luckmann's notion of reality system.[2] We can thus meaningfully draw upon Kuhn's ideas in building a historical dimension into our model. Kuhn's work has additional relevance in that it focusses upon 'specalised' knowledge which is our particular concern. Paradigms are shown to play a crucial part in the differential patterning of the knowledge generation process over time.

In an early formulation of his theory, Kuhn distinguished between preparadigmatic and paradigmatic stages in the development of a field. His argument was that, in the preparadigmatic stage, fact-gathering is random (cf Plinean, Baconian natural histories). Then paradigms emerge to guide fact-gathering, one of which may triumph over its competitors. In the postparadigm period of its life, a field goes through cycles in which periods of 'normal' paradigm-guided activity are punctuated by periods of crisis, when recognition of weakness in the current paradigm coincides with the emergence of a competing paradigm or paradigms. After a 'revolution', in which one of these new paradigms overthrows the old one, the field settles down to a further period of 'normal' activity until the next point of crisis occurs.

---

[1] Kuhn, T.S. Op. cit.

[2] A recent description of what Kuhn intends us to understand by 'paradigm' brings it closer still to the notion of 'reality system'. Kuhn argues that 'languages cut up the world in different ways'. He goes on to ask how 'we acquire the knowledge of nature that is built into language'. The answer is that we do so 'by the same techniques and at the same time as we acquire language itself, whether it is everyday or scientific', a process which is 'not fully linguistic'. Learned relationships are deployed unproblematically, 'yet without being able to name the characteristics by which we make the identifications and discriminations'. In other words: 'They are prior ... to a list of criteria which, joined in a symbolic generalization, would enable us to define our terms. Rather they are part of a language-conditioned or language-correlated way of seeing the world. Until we have acquired them, we do not see a world at all.' (Kuhn, T.S. Reflections on my critics, in Lakatos, I. and Musgrave, A. eds. Op. cit.)

Thinking at this level, we can see that in seeking to reflect change in our system, providing we are dealing with a paradigm-guided field, we are not faced with random variation over time. Variation is patterned in terms of paradigm change. The activities of paradigm holders have predictability. The differences which occur in association with paradigm change are the basis on which we may work.

It is necessary to consider Kuhn's theory in more detail to explain why it is essential to mirror change in our system. Kuhn argues that a paradigm identifies the problems that can be assumed to have solutions, and the techniques, instruments of measurement and so on which are appropriate to their investigation. It also provides the rules (embedded in 'exemplary' problems, and intuitively understood rather than explicitly formulated) which define acceptable solutions. As 'facts' come to light which do not fit the paradigm, like jigsaw pieces belonging to another puzzle, 'extraordinary' science takes over from 'normal' activity (in science this is 'puzzle-solving'). At such a time the rules themselves are tested for weaknesses and research, in a problem-directed sense, becomes random. Alternative theories emerge and ultimately, by what Kuhn can only describe as a 'Gestalt switch', a new paradigm becomes accepted. The new paradigm focusses upon a new range of problems and provides new rules defining acceptable solutions.

It is crucial to understand the effect of this paradigm change. The scientist's world changes in this process, and even the entities of which his world is composed are redefined. The same kind of activity is resumed, but the scientist now sees different things when looking in the same places with the same instruments. If we aim to provide stability in the sense of preserving the entities within our system, in the face of paradigm change, it will soon reflect a world of entities which no longer 'exists', and will not provide categories for documents concerning the world as currently perceived. The user will then no longer be able to function in a 'normal' manner.[1] This contrasts with a view which holds that

---

[1] The distinction between 'normal' and 'extraordinary' science, on which the notion of 'Gestalt switch' depends, has been questioned. We tend to agree with Watkins that the *individual* scientist, given an accepted paradigm, is nevertheless able 'to consider it critically, or to toy with (without necessarily embracing) alternatives to it'. This means that puzzle-solving is not the principal activity and that 'the scientific community is not, after all, a closed society whose chief characteristic is "the abandonment of critical discourse"'. (Watkins, J. Against 'normal science', in Lakatos, I. and Musgrave, A. eds. *Criticism and the growth of knowledge*. (Proceedings of the International Colloquium in the Philosophy of Science, London, 1965. Vol.4.) Cambridge, Cambridge University Press, 1970.) Considering scientific activity in general, Toulmin challenges the argument that 'the differences between the kinds of change taking place during "normal" and "revolutionary" phases of scientific development are, at the intellectual level, absolute'. (Toulmin, S. Does the distinction between normal and revolutionary science hold water?, in Lakatos, I. and Musgrave, A. eds. *Op. cit*.) However, whether the difference is one of degree or kind, even such critics seem to agree with Kuhn that the development of science is characterised *at the group level* by 'conceptual transformations'; the main point at issue is rather whether these are in some sense the 'essence' of science or rarer in occurrence and to be 'downvalued', by contrast with 'normal' science, to something analogous to the 'confusion and despair' which accompanies a religious conversion. Toulmin thus argues that 'No theory of scientific growth and development would be adequate which did not recognise, and do justice to these intellectual discontinuities'. He suggests that 'One may question, indeed, whether any natural science having a serious theoretical component *ever* develops by a process of "accretion" alone'. Masterman argues, unlike Watkins, that 'normal' (ie puzzle-solving), not 'extraordinary' (ie critical discourse) activity, is the 'essence' of science. Like Kuhn, she regards a paradigm as a 'way of seeing', although she points out various problems with regard to Kuhn's arguments. Again like Toulmin, however, she believes that, even if we reject Kuhn's conclusions: 'we are not going to be able to go back to where we were before Kuhn and his immediate predecessors began to get at us. Their protest against the unconscious dishonesty ... with which the history of science has been done in scientific textbooks up to now cuts far too deep; and so does their outcry against the oversimple and distorted accumulative view of science which has resulted from reading the textbooks as though they were real history.' (Masterman, M. The nature of a paradigm, in Lakatos, I. and

continued...

stability of the kind we have described is essential for systematic activity and, in the information context, a prerequisite for retrospective searching. This fails to take account of the discontinuity which may occur in the process of paradigm change and its relation with the knowledge generation process.

We noted, however, that it is practically viable to mirror change in a system only if activity in a field is paradigm-guided and hence change is patterned. We must therefore consider the status of the social sciences in relation to the categories in Kuhn's theory. Kuhn initially suggested that, in the terms of his theory, the social sciences were at the preparadigmatic stage of development. Hence random change, deriving from lack of systematic activity, would be expected. Crane[1] describes sociometric and bibliometric studies revealing an apparent lack of social organisation in the social sciences as contrasted with the sciences, which appear to support this assumption. If we study the knowledge generation process itself, however, we find a situation which is far from one of

---

[1] Crane's survey covers recent work in these areas. (Crane, D. Information needs and uses, in Cuadra, C.A. ed. Op. cit.)

---

continued

Musgrave, A. eds. Op. cit.) It is not necessary for us to enter into the controversy surrounding the notion of 'Gestalt switch' in which critics have engaged, nor that concerning ideology v rationality as explanatory principles of change. We do not need to take sides over the nature of the relation between adherence to a paradigm and exploration of alternative paradigms, but can regard this as a matter for further investigation. Whether scientific activity is best characterised as critical discourse, or puzzle-solving, or some combination of the two, and whether fields are uniform or differ in this respect, there is agreement that such activity is paradigm-related. The general point of relevance for our model which emerges from the discussion is the recognition that this activity is marked by 'intellectual discontinuities', and 'conceptual transformation'; this is a phenomenon central in some way to the theories of change of all these writers.

unsystematic fact gathering. Activity is paradigm-guided[1] and thus, by definition, patterned. This difference in assessment of the situation appears to be due mainly to problems inherent in Kuhn's theory, to which Kuhn has since drawn attention. He has also stressed that his theory is based upon study of the sciences; it is for others to examine its bearing upon other fields.

As outlined so far, the theory relates to an abstraction from the knowledge generation process over time in terms of a single community of people, defined as a community because they share a paradigm, and has looked at the life course of their paradigm. 'Normal' activity has been associated with a single paradigm. When we look across the spectrum of subject fields, we note that whereas science fields tend to be characterised by consensus upon a single paradigm, social science fields lack such consensus. However, this does not mean that 'normal' (ie paradigm-guided) activity is not to be found in social science fields. Kuhn points out a circularity inherent in the notion of paradigm. It is as true to say that 'communities have paradigms' as that 'having a paradigm' defines a community. Whether a community is found to have one or many paradigms depends upon the unit of analysis (eg specialty v discipline). Masterman[2] believes we should distinguish a state of affairs termed 'multiple-paradigm science': '... multiple-paradigm science is full science, on Kuhn's own criteria; with the proviso that these criteria have to be applied by treating each sub-field as a separate field.'

---

[1] Work in our field also tends to cluster at any time around a particular problem, until the focus shifts elsewhere. Young describes how in England over the last fifteen to twenty years 'the foci have been *equality of opportunity* and the *wastage of talent*, *organization* and *selection of pupils*, and the *curriculum*'. (Young, M.F.D. An approach to the study of curricula as socially organized knowledge, in Young, M.F.D. ed. *Op. cit*.) Such foci seem to resemble the phenomenon Diana Crane terms 'fashion'. (Crane, D. Fashion in science: does it exist? *Social Problems*, 1969, 16(4), 433-441.) This phenomenon cuts across normal/crisis periods. Crane suggests that popular areas may have 'strong theoretical imperatives' attracting scholars during a period of normal science, though in prenormal and postnormal periods (ie just after the establishment of a paradigm when the area is not 'visible', and when its immediate implications have been exhausted and the problems are difficult) its attraction may not be so great. Equally, however, she finds clustering where no strong theoretical justification exists. Masterman refers to 'non-paradigm science'. (Masterman, M. *Op. cit*.) However, to a large extent, focus on a particular range of problems seems to be paradigm-bound in our field. Bernstein, for instance, notes a shift from the structural functionalism of the social class and educational achievement 'period' to phenomenological studies of the curriculum. (Bernstein, B. *Op. cit*.)

[2] Masterman, M. *Op. cit*.

This fits well with Kuhn's suggestion that 'communities can and should be isolated without prior recourse to paradigms; the latter can then be discovered by scrutinising the behavior of a given community's members'. It avoids the danger of confounding pre-paradigmatic, crisis and multiparadigmatic states, in which a common factor is that all lack a single paradigm. This is a danger even in the science context. Kuhn notes that 'There are schools in the sciences, communities that is, which approach the same subject from incompatible viewpoints'. Lodahl and Gordon[1] found a lack of consensus in the physics field which they attributed to a crisis state but which seemed to have much in common with the situation in sociology, hypothesised to lack well developed paradigms. More importantly, from the social scientist's point of view, it leaves open the question of whether the state to which a social science is or should be tending is the single paradigm state.[2] It seems quite possible that a situation of coexisting paradigms may be a necessary condition for healthy social science.

Returning now to our model, we characterise our field as multi-paradigmatic, and cognitive sub-universes are the 'communities' in which we are interested. Each paradigm has its individual life course and associated 'normal' activity. (For this reason, each must be handled separately in our system.) Looking at the field as a whole at any point in time, we see apparent disorganisation because different paradigms are at different stages in their life courses. At any series of later points in time there will be a similar impression of disorganisation, although the picture will on each occasion be different, suggesting random variation but in fact showing our paradigms at successive points in their various life courses. From Kuhn's historical perspective there is systematic patterning over time, the complexity of which may tend to obscure it from the observer at any given time.

---

[1] Lodahl, J.B. and Gordon, G. The structure of scientific fields and the functioning of university graduate departments. American Sociological Review, 1972, 37(2), 57-72.

[2] Cf Feyerabend's questioning whether, as a matter either of description or prescription, the character of science is to be viewed as being 'as monolithic as Kuhn makes it out to be'. (Feyerabend, P. Consolations for the specialist, in Lakatos, I. and Musgrave, A. eds. Op. cit.)

In considering the processes underlying this patterning, Kuhn suggests that, once a 'way of seeing' is acquired, objects and situations are grouped 'without an answer to the question, "similar with respect to what?"' Paradigm change occurs when 'Objects which were grouped in the same set before are grouped in different sets afterwards and vice versa'. It is in this process that entities are redefined; 'alloys were compounds before Dalton, mixtures after'. Whilst 'the names of the sets are generally preserved', nevertheless 'the transfer of a subset can crucially affect the network of interrelations among sets'. Differences of this sort are 'not simply about names ... but ... about nature'.

For a more detailed study of the dynamics of change we may look to some concepts in a model proposed by Smucker and Zijderveld.[1] Social institutions, in which we may include socially constructed reality systems and the languages in which they are embedded, are viewed as 'meaning-structures', and their dialectical nature is stressed. Meaning and structure are kept conceptually separate for the purposes of discussion, although at any point in time they coalesce.

The writers distinguish between change as a consequence of an internal dialectic and change as a consequence of the transferral of elements from one institution to another. Both meaning and structural components are subject to both types of change, and there may be interaction between them. Change is endemic in all institutions, and change of meaning in a term as it is used within one reality system may alter the meaning of others used within the system, and possibly the structure of relations amongst terms. Transferral of meaning may occur when one reality system finds 'an appropriate model of explanation' in another. This can produce change of either meaning or structure or both in both systems. In turn the changes will trigger off others.

---

[1] Smucker, M.J. and Zijderveld, A.C. Structure and meaning: implications for the analysis of social change. British Journal of Sociology, 1970, XXI(4), 375-389.

Change may be disruptive or non-disruptive. For instance, empirical investigation may show that a given term, as used within a given reality system, is no longer 'legitimised by a given content of meaning'. There are two possible effects. There may be structural change within the institution but meaning will be maintained. Alternatively the structural relations may be maintained but new meanings attached to the elements. Thus one component of an institution adapts to the other, or in the case of transferral, if one reality system takes over and redefines a concept from another, there will be adaptation of boundaries between institutions, in terms either of structure or meaning. When adaptive mechanisms fail however, there is likely to be disruptive change and schism, in which total new reality systems will emerge and existing systems will possibly disappear or be redefined. The higher the degree of differentiation within the society, the greater the rate of change, simply because change is endemic and there is more that can change.

This model also shows clearly that mechanisms for change must be built in if our system is not rapidly (because of high overall differentiation we may expect a considerable degree of change) to become out of touch with thinking in the field, and hence to be ill-adapted for the purpose of knowledge generation. We hypothesise that the rapidity of change will be proportional to the extent to which, because knowledge in a reality system is permeable (ie not accepted as objectively real), communication depends upon shared understandings.

Some change will have a greater bearing on the effectiveness of knowledge generation than other change. The distinction between adaptive and disruptive change (which we may loosely associate with normal and crisis states) is helpful here. Adaptive change in a 'normal' situation may be considered in relation to degree of permeability within and across reality systems. Within high permeability reality systems, the shared understandings which users of such systems bring to them will supply the necessary adjustments. In reality systems with low permeability, adaptive change will be rare because their knowledge is regarded as objectively real. We hypothesise therefore that, in both cases, adaptive change is unlikely seriously to impair the knowledge generation process. Thus our model need take account of adaptive change only from a long term point of view.

We may now consider disruptive change, and relate it also to degree of permeability. Disruptive change will involve the decay of total reality systems and the emergence of total new ones, though clearly this will not happen overnight. Such change will produce further change in other reality systems not directly involved in 'extraordinary' activity, but this will probably be of an adaptive nature. Some portion of the 'landscape' will, however, be completely changed, and we may expect 'objective' and 'subjective' systems to be equally susceptible. We hypothesise that effective knowledge generation would rapidly become impossible if account were not taken of disruptive change in the patterning of reality systems, because it concerns not minor adjustments of meaning but changes the bases upon which meanings are imputed. Our model must therefore allow for the notions of discontinuity and transformation.

(d) Theories of semantics*

The ideas we have discussed so far have enabled us to specify the elements of our model in formal terms. We have been able to consider general characteristics of the knowledge generation process and of reality systems. In order to proceed further, and to specify the intellectual dimensions of knowledge generation, we may turn to theories of semantics since, as Berger and Luckmann point out, language is the medium for the social construction of reality, and reality systems are embedded in the language we use. There will be implications in terms both of intellectual organisation for our system, and of the language with which we equip it.

---

* We are extremely grateful to David Stringer, of the Faculty of Educational Studies, Open University, for reading and discussing drafts of this section of our paper.

It is helpful to distinguish between philosophical and linguistic semantics. Lyons[1] gives a useful general survey of the principles on which linguists are working. He comments that philosophers have always been interested in meaning, but 'we soon find ourselves in the thick of philosophical controversy'. The problem is to clarify what is meant by meaning! One central debate concerns the nominalist/realist issue: 'Have the things to which we apply the same name some common "essential" properties by which we can identify them (as the "realist" might say) or have they nothing in common other than the name that by convention we have learned to apply to them (as the "nominalist" might say)?' Abstract notions also present serious problems. 'Do all the things we describe as "beautiful" or "good" have some common property?' We may say that 'beauty' or 'goodness' is a concept associated in the minds of those who use the terms with things thus described. But there is a further debate over the appropriateness of thinking of concepts as 'existing'.

The notion of reference bypasses these problems: a word is held to be related to a referent 'through the mediating (conceptual) meaning associated with both independently'. Traditional categories of sameness, difference and relatedness may then be employed. But this will not be helpful with terms such as ours which often have no direct empirical referent. And, as Lyons reminds us, 'there is no one point in the vocabulary from which you can start and from which you can derive the meaning of the rest'.

Frequently a word can only be understood by 'putting it in a context' (cf Berger and Luckmann's 'shared understandings'). Lyons quotes Wittgenstein's maxim 'Don't look for the meaning of a word, look for its use'. This draws attention to the fallacy of assuming that words necessarily have fully determined meanings. For Lyons, 'the way in which language is used in normal situations can be explained on the much weaker assumption that the speakers of the language in question are in sufficient agreement about the use of words ... to prevent "misunderstandings"'. He further suggests that it is undesirable that words should have fully determined meanings. Once we take account of context, it is clear that fully determined meanings can only inhibit communication. In communication, 'the context is constantly developing, in the sense that it "takes into itself" ... all that is relevant to the production and understanding of further utterances'. Even thinking in terms of reference, 'some philosophers have suggested, referential "impreciseness" ... makes language a more efficient means of communication'. Lyons emphasises, however, that no-one is 'in a position to make any very definite statements about the way in which a knowledge of the abstract relationships holding between grammatical elements in sentences interacts with contextual features of various kinds ...'.

---

[1] Lyons, J. Introduction to theoretical linguistics. Cambridge, Cambridge University Press, 1968.

Linguists have generally proceeded on the basis of a methodological decision to assume a 'restricted context', as Lyons points out, so that it is then possible at least to study 'sense': 'By the sense of a word we mean its place in a system of relationships which it contracts with other words in the vocabulary'. This shelves the problem of the relationship between words and the world. The approach is through structuralism. Lyons defines this approach as one in which 'each language is regarded as a system of relations (more precisely a set of inter-related systems), the elements of which - sounds, words etc. - have no validity independently of the relations of equivalence and contrast which hold between them'. This is a formal mode of analysis, in which the object of analysis is the structure rather than the content of its elements.

The assumptions entailed by this approach are made very clear by Chomsky[1]: 'Linguistic theory is concerned primarily with an ideal speaker-listener, in a completely homogeneous speech community, who knows its language perfectly and is unaffected by such grammatically irrelevant conditions as memory limitations, distractions, shifts of attention and interest, and errors (random or characteristic) in applying his knowledge of the language in actual performance. This seems to me to have been the position of the founders of modern general linguistics, and no cogent reason for modifying it has been offered. To study actual linguistic performance we must consider the interaction of a variety of factors, of which the underlying competence of the speaker-hearer is only one. In this respect, the study of language is no different from the empirical investigation of other complex phenomena.

---

[1] Chomsky, N. Aspects of the theory of syntax. The Hague and Paris, Mouton, 1965.

'We thus make a fundamental distinction between competence (the speaker-hearer's knowledge of his language) and performance (the actual use of language in concrete situations). Only under the idealisation set forth in the preceding paragraph is performance a direct reflection of competence.'

We established earlier the potential usefulness of structuralist concepts for us. But we find that structuralism, as it is employed within linguistic analysis, entails assumptions which, whilst affording valuable insights to the linguist, are in principle unacceptable for the situation we are investigating.

First, the assumption of the 'ideal speaker-hearer' is seriously inappropriate in a speech community such as our users represent. The differences in the 'languages' spoken by our users are essential to effective communication.

Secondly, Chomsky stresses that linguistic theory focusses upon 'competence'. This is an abstraction, in contrast to 'performance', which refers to use of language. We are concerned not with abstraction from language but actual use of language, so as to distinguish different reality systems in terms of the various 'languages' in which they are embedded, and also to label and describe appropriately the documents 'belonging' to them.

Thirdly, we must bear in mind the problematic nature of the relation between syntactic and semantic aspects of language. The early Chomsky assumed that formal syntactic analysis could proceed without consideration of semantic problems, and made the now well known distinction between deep and surface structure. He later argued that meaning was closely associated with deep structure and could be subjected to the same kind of analysis as syntax. However there are considerable difficulties in divorcing semantics from syntax. Lyons, for instance, recommends a degree of scepticism, at the present stage of understanding, before regarding surface structure as irrelevant to semantic interpretation. He stresses that 'no-one has yet presented even the outlines of a satisfactory and comprehensive theory of semantics'. In particular he questions the assumption that semantic components are universal.

A study carried out by Conklin[1] in the Philippines is relevant to this latter question. He shows that syntactic structure is not isomorphic with semantic structure. The rules governing semantic relations amongst categories in lexical sets cannot be prescribed merely on the basis of 'familiarity in another system with the "concrete" denotata of the sets involved'. He instances local systems of plant classifications. This is so despite the fact that the lexical items represent part of the everyday vocabulary of the language. Multiple dimensions may be involved as in the case of a set of eight pronouns that seemed to have no logic until they were found to require three dimensions to define them: minimal/non-minimal membership; inclusion/exclusion of hearer; inclusion/exclusion of speaker. Such a set is ordered by class inter-section and cannot adequately be represented as hierarchy, or a circular, block or branching diagram. This seems to justify Lyons in his caution.

Frake[2] emphasises one cannot study the language of a community without understanding of its culture. He proposes that, instead of 'getting words for things', the ethnographer should find 'the things that go with the words', or 'finding out what are in fact the "things" in the environment of the people being studied'. We need to discern 'how people construe their world of experience from the way they talk about it', or in other words concern ourselves with 'the analysis of terminological systems in a way which reveals the conceptual principles that generate them'. The central point is that 'Even with reference to quite obvious kinds of material objects, it has long been noted that many people do not see "things" quite the way we do'. This has often been dismissed as 'primitive thinking'. In fact there is no reason why one group of people should select as significant the same dimensions of an object or a notion as another group, and hence classify them in the same way. What is significant is defined by local cultural factors.

---

[1] Conklin, H.C. Lexicographical treatment of folk taxonomies, in Fishman, J.A. ed. Readings in the sociology of language. The Hague and Paris, Mouton, 1968.

[2] Frake, C.O. The ethnographic study of cognitive systems, in Fishman, J.A. ed. Op. cit.

In general, if one is interested in the use of language as a phenomenon, one may take issue with linguistic analysis over the non-social abstraction approach employed, even whilst accepting Chomsky's point that underlying competence is one of the variables to be considered.

This rather dismal conclusion seems to be endorsed by Kay and Sparck Jones.[1] Their assessment of the situation is that 'A large proportion of today's linguists ... are concerned with these problems of so-called "linguistic universals".' Progress is being made in empirical work in 'linguistic engineering', but it is difficult to predict 'how the ferment of theoretical linguistics will affect the progress of linguistic engineering'. The writers note that 'a disproportionate amount of effort in computational linguistics is devoted to syntax', although 'interest in non-trivial treatments of semantic problems by computer is growing'.

There is endorsement also from Hymes,[2] who is critical that 'Sometimes as a matter of simplifying assumption, sometimes as a matter of principle, linguistic theory has been almost exclusively concerned with the nature of a single homogeneous code, shared by a single homogeneous community of users, and (by implication) used in a single function, that of referential statement'. Bilingual communities are obvious cases where this assumption does not hold. Yet these are simply 'special cases of the general phenomena of variety in code repertoire and switching among codes'. However, 'there is not even agreement on a mode of descriptive analysis'.

---

[1] Kay, M. and Sparck Jones, K. Automated language processing, in Cuadra, C.A. ed. (Op. cit.)

[2] Hymes, D. Models of the interaction of language and social setting. Journal of Social Issues, 1967, XXIII(2), 8-28.

Sociolinguistics is a relatively new field which Hymes regards as a 'movement to redress the situation'. Unfortunately, since even the procedures for studying non-homogeneous communities such as ours have hardly begun to be developed as yet, there is no guidance as to the conceptual principles we may expect to find. At this stage there seems to be no option[1] therefore but to proceed informally in an exploratory way, working with users on their material, in seeking to understand their 'culture' in the way suggested by Frake.

Despite the fact that we must start 'from cold', as it were, in specifying the intellectual dimensions on which knowledge becomes structured in our field, the insights we have gained into the principles upon which mainstream linguists work are of considerable relevance in considering the formal characteristics of an indexing language.

It seems clear that our users had some degree of reason on their side in their resistence to formal control of terminology in relation to material in their field, and the imposition of

---

[1] Content analysis might seem to represent an alternative approach, although Kay and Sparck Jones (Op. cit.) have reservations about this type of technique. Krippendorff suggested that content analysis 'can be considered as the operationalization of a kind of semantic theory that could either be put to test in experimental situations involving human subjects or, more efficiently, be deduced from a more general theory if it were available'. (Krippendorff, K. Theories and analytical constructs, in Gerbner, G. ed. The analysis of communication content: development in scientific theories and computer techniques. New York and London, John Wiley, 1969.) Unfortunately, Krippendorff finds that 'more is known about the formal structures of theories of symbolic behavior than about their actual content'. For this reason much of content analysis is 'justified on entirely intuitive grounds'.

'syntax' or structure. It may well be different when our principal interest in words is as referents to empirical objects. Once the importance of context for communication in a field is established, however, the fact that the relation between features of language and contextual features is not fully understood, even by linguists, suggests that we are highly likely to destroy important cues and cause information loss. This seems to hold whether we think of the elimination of synonyms, the choice and consistent use of a preferred citation order, the construction of hierarchies, or other aspects of control.

The reason is that such procedures may cut across 'natural' structures. We find that even relations such as the genus/species one are not universal (cf Conklin's study of a pronomial system). Something very close to Chomsky's notions of 'ideal speaker-hearer' and 'competence' appears to underlie the notion of control in indexing languages. The alternative seems to be the unpredictability of natural language, which is inefficient for 'retrieval' purposes. Yet sociolinguists are beginning to find regularities in 'surface' variation, in 'performance'. These regularities are obliterated if we base our systems on something analogous to 'deep structure', and yet it is at the level of 'surface structure' that users interact with documents and are able to function by means of shared understandings.

Such regularities represent an intermediate level at which to structure our indexing language, one which corresponds roughly with that of the cognitive universe which we have taken as our unit of study. Such an approach rejects for practical purposes, at least at the present time, the assumption of 'universals', along with the idealisation of the actual language situation. It avoids the practical disadvantages of a so-called unstructured natural language. It looks instead for the particular structure of relations between terms (our concept of permeability is relevant here) as used within each cognitive subuniverse.

Part II Methodology: exploratory study in one social science field.

2. The construction of a system

Nature of guiding model

We may briefly summarise the model we have outlined as a characterisation of the social nature of the knowledge generation process, and the part to be played in it by an information system. Leaving aside idiosyncracies of an individual's meaning system, users are seen to form cognitive subuniverses in terms of a background of shared understandings with significant others, shared symbol system etc. or, in sum, a reality system. The nature of the reality system will determine jointly the user's mode of interaction with others and the way in which knowledge is seen to be structured. Mode of interaction may be characterised in terms of the extent to which shared understandings play a necessary part in knowledge generation; the structural aspect of the process may be considered in terms of the permeability of knowledge within and across reality systems. Additional factors are the processes associated with the emergence and decay of reality systems over time. An effective information system must allow maximum opportunity for all users to function as members of cognitive subuniverses. If we can devise such a system we shall have taken a big step towards the ideal of an objective (ie non-interfering) system.

Functioning of a system

We may describe the user as wishing to locate significant others in the literature, ie those with whom he shares common understandings. This is the means by which he will identify useful documents. A sharing of understandings is associated with a shared symbol and reality system. (A user may of course be at home with more than one perspective or reality system.) The user will interact in different ways with 'others' in the literature. His starting point will be determined by the perspective he is using at the time. His mode of interaction with the 'others' he meets there will be determined by the extent to which their perspective is compatible with his own reality system (ie they are for him significant others).

Distinctive reality systems should be handled separately by the system, and internal structure will be individual to the reality system. But differing reality systems will not necessarily be incompatible. They may represent simply different versions of the same reality. It will not always be easy to determine which is the case, and if so the options must be left open to the user. There will be varying degrees of permeability amongst reality systems. The permeability of any reality system will not be invariant but a function of its particular relation with other individual systems. This will have consequences for the meaning of terms in our system. Also independent of the question of permeability but complicating it, the same term may have different meanings in different reality systems. There is thus a threshold to be passed if the user is to use the system in a way which takes account of these problems. The effectiveness with which he is able to use the system will depend upon the background (or context) he brings to his use of the index. At the same time the system must be designed to provide contextually relevant data on which to base the distinctions his background equips him to make.

In operating within the system from a chosen entry point, the user will 'translate' the ideas in the literature he has found in terms of his initial perspective, and will accept or reject them, as they are then meaningful to him, in relation to his problem. He may then withdraw satisfied, or may decide to redefine his problem and reenter from another perspective. Alternatively he may choose to translate his ideas into those of the various reality systems he meets and is familiar with (ie he speaks the 'language') in the range of literature he encounters. On this basis he may select a perspective appropriate to his problem and proceed further on this new basis. This switching of perspective may occur again later as he responds to the 'others' he meets as a result of the path he makes for himself through the literature.

Overall framework of the system

We have decided that reality systems should provide the general framework of our system. Central to this is the notion of 'shared understandings'. Certain aspects of this notion, such as role (eg teaching/research), field of study and other variables, have been generally recognised as relevant. One additional and, we believe, major variable particularly relevant to our field is intellectual perspective or conceptual model. (This is not to say that this variable is not relevant to other fields and is not recognised implicitly in, say, cross-disciplinary studies of the hard sciences.)

No simple distinction in discipline terms will serve our purpose. Hansen[1] has discussed differences in perspective between sociologists and educators. A particular problem this raises for the sociological analysis of education is the consequence for sociological concepts. Kurtz[2] discusses 'the adoption and misuse' of sociological concepts, with particular reference to 'culture' and 'social class'. For the sociologist, a culture represents a widely held pattern of 'values, objectives, norms, and knowledges', the pattern in the subculture differing from that in the more general culture. Cultural deprivation, therefore, is strictly a term of limited application as, for instance, to describe extreme social isolation. But in lay use it has come to be equated with 'social groupings which are on the bottom rung of the ... socioeconomic ladder'. This leads to further confounding by the association of mental retardation with the lower class. Lay usage is accepted usage, but at the same time the lack of a technical vocabulary reserved as such 'may be disabling' to the sociologist.[3] There is the converse of this problem also to be considered: lay terms adopted by the sociologist which are never wholly divested of everyday connections.

We have also to consider differences in perspective amongst the different disciplines brought to bear upon the study of education.[1] Differences are to be found both as between sociologists and non-sociologists studying education, and within each of these groups. Morris,[2] writing on psychology and education, contrasts an adult-centred view of education involving 'doing things to people' with a child-centred one 'concerned with individual experience and the development of powers'. The former reflects a view of the individual simply as an organism responding to stimuli; 'the behaviour of an organism is as far as possible to be reduced to a mechanistic model'. The 'mode of thought' to which the latter 'is most easily assimilated is what we may call an organic one'. Morris notes that common to both these views is 'a rigid separation of society and the individual'. This is antithetical to sociology. Comparative educationalists often, in discussing education institutions, suggest considerations (eg of social organisation) possibly far closer to those in some sociological writing.

---

[1] Hansen, D.A. The uncomfortable relation of sociology and education, in Hansen, D.A. and Gerstl, J.E. On education: sociological perspectives. New York etc, John Wiley, 1967.

[2] Kurtz, R.A. The public use of sociological concepts: culture and social class. American Sociologist, 1966, 1(4), 187-189.

[3] Mortimore, G. and Enfield, R. Op. cit.

It is even more true to say that one cannot generalise about sociology. Dawe[3] notes, for instance, a fundamental difference of perspective as between those sociologists for whom 'order' and 'control' respectively represent their major organising concept. Such differences can be cross-cutting to the point at which, as Dawe illustrates, concepts associated with one of the perspectives 'disappear' when subsumed under concepts associated with another. In the perspectives he discusses, 'the point is that, as soon as definitions of the situation become properties of the central value system - that is, as soon as the elements of action are, in effect, reduced to the single element of situational conditions - then ... action disappears.'

---

[1] From Dershimer's summary of a colloquium about educational research, it seems that although invisible colleges may exist within the field, researchers tend to 'turn to discipline-based rather than education-based colleagues for information, critical review and approbation'. The distinction between 'invisible college' and 'reference group' is by no means a clear one. (Dershimer, R.A. The educational research community: its communication and social structure. Washington, D.C., U.S. Department of Health, Education and Welfare, 1970.)

[2] Morris, B. The contribution of psychology to the study of education, in Tibble, J.W. ed. The study of education. London, Routledge and Kegan Paul, 1966.

[3] Dawe, A. Op. cit.

Another problem is the way in which concepts are used. We have earlier noted a distinction between analytical and sensitising concepts. Even in the case of relatively analytical concepts, distinctions amongst 'achievement', 'aptitude', 'intelligence', for instance, are seen to be meaningless when it is realised that performance on tests designed to measure any one of these may be taken as an indicator of any other, and the results viewed as contributing to understanding of the nature of this other as a construct. Additionally, in the light of empirical work, concepts are constantly being redefined better to explain the 'facts'. Formal or so-called fundamental categories such as genus/species and even facet-like categories such as structure/process raise many problems for the classifier in our field.[1] For sociologists, for instance, structure is process viewed at a single point in time, and any document must be classified so as to leave the user free to take either a structural or processual view. We need to be sensitive to the ways our users think.[2]

In the sociology of education project, therefore, we are specifying the reality systems which must feature in our system by investigation of the range of accepted ways in which authors, in their documents, categorise work in the field. We find that authors variously associate themselves with others[1] in terms of substantive problem area, a particular concept or combination of concepts, conceptual model, 'tradition' (eg Weberian, phenomenological), and so on. Broad classes of theoretical explanation (eg subcultures) or analytical dimensions (eg local/cosmopolitan) are current at any given time. These range up to general perspectives at the level of Dawe's order/control distinction. Our evidence comes from analysis of the literature in consultation with authors, and from discussion on use-strategy with users.

---

[1] The return to disciplines as main classes in Jean Aitchison's thesaurofacet recognises this point. (Aitchison, J. et al. Thesaurofacet: a thesaurus and faceted classification for engineering and related subjects. Whetstone, English Electric Co., 1970.)

[2] It is interesting to find Crane commenting upon 'the variety of ways in which terms can be used in scientific documents'. She believes that 'further progress in the development of devices of all sorts to speed information-seeking will come about as a result of basic research on how scientists use ideas, which is in turn related to understanding how knowledge develops in an area ...'. (Crane, D. Information needs and uses, in Cuadra, C.A. ed. Op. cit.)

Authors are also users, and may be expected to influence the way in which both author and non-author colleagues view work in the field. If it is reasonable to suppose that users 'take over' the categories of authors as significant others, and that they will be oriented by them in approaching the literature, then the literature may helpfully be organised correspondingly, and we have a system of 'labels' employed within the field for the purpose.[2]

The different kinds of categorisations we are building into our system are cross-cutting, not complementary. Consequently in the overall structure of our scheme for intellectual organisation, we are handling the different aspects of intellectual structure, or ways of thinking about work in the subject field, which we have identified, as separate frameworks.[1] We assume that whilst the user may 'cut the cake' in various ways, he will not cut it in more than one way at a time. We are now at the point of experimenting to discover how we may most usefully (relative to the particular type of user 'strategy' we associate with a given perspective) structure each framework internally.[2]

---

[1] In identifying the specific categories which should have a place in our scheme, we find that there are commonly explicit statements in the text about the different ways in which an author characterises his work in relation to that of others. Where this is not so, it may be possible to induce the labels he would apply to himself from the nature of his work, but there is often ambiguity. There is a similar problem if we try to categorise his work in terms of the way he describes his significant others. Lipetz has pointed out that lists of references alone are not helpful because they lack explanation of the relation between citing and cited item, ie why work has been cited. (Lipetz, B. Improvement of the selectivity of citation indexes to science literature through inclusion of citation relationship indicators. American Documentation, 1965, 16(2), 81-90.) There is then ambiguity as to the categories underlying items selected for reference. In our case the author's 'labels' and his personal standpoint, when unclear, are being clarified by asking him. We have found authors very ready to help when there is a risk of possible misrepresentation of their thinking.

[2] This should be helpful in avoiding the situation, discussed by Crane (Op. cit.) in which 'The labels that scientists assign to [research areas] do not necessarily correspond to the categories used by journals or indexing and abstracting services'.

The centrality of the notion of permeability now becomes clearer. At the level of different versions of the same reality, there will be little distortion of meaning if distinctions are blurred because the individual may supply the necessary adjustments for understanding. Where different reality systems are concerned, such as those Dawe refers to, even participation on the part of the user cannot restore concepts which have 'disappeared'.[3] Within frameworks, where there is permeability, concepts may be 'merged' as headings without substantially modifying the presentation of the literature as a corpus. We would go further and argue that they should be merged, since permeability means that distinctions will be arbitrary and hence unhelpful. Thus, achievement, intelligence, aptitude and other related terms will be treated as a single concept cluster, because of the way in which, in use, these notions are intertwined, as we noted earlier.

The reason for stressing that the structure of ideas is inseparable (except in an analytical sense) from operations or 'search strategy' also becomes clearer. These elements jointly bear, via the concept of permeability, upon the desirable organisation of a store of documents as represented by an information tool.

---

[1] Framework in this context should not be taken to denote 'facet' in another guise. Frameworks here are to be understood to define and present the same universe in different ways, hence offering alternative and cross-cutting total representations of the same universe, and not alternative but partial and complementary subsections of the same universe. A closer analogy is with the alternative approaches to the contents of a bibliography offered by separate author, title and subject indexes.

[2] Our scheme is still tentative but details can be made available to interested enquirers.

[3] This is a principle which to some extent appears to be implicit in the structure of Aitchison's 'Thesaurofacet' (Op. cit.). The terms in which a given area of subject matter is described are not necessarily identical in alphabetical and classified sections of the tool.

Where consensus is high, the existence of boundaries between reality systems may be expected to be of little concern to the user, but clear boundaries within the system are crucial. Meanings will be relatively fully determined and meaningfulness established largely on the basis of individual terms without the need for further insights from shared understandings. A simple matching process is likely to be the obvious mode of search. Where consensus is low, sharp boundaries within the reality system are undesirable but the maintenance of boundaries between a given system and others are crucial. The classification should be sensitive to these considerations and, in cases where there can be no clear cut decision as to level of permeability, the options must be left open for the user. The mode of search will in this case necessarily be open ended if the search is to be at all effective. Use of the system will in fact be an act of knowledge generation, and as such the required outcome cannot be specified in advance.

The representation of individual documents in the system

Just as overall structure of our system must be meshed with different modes of 'search' or interaction with a corpus of documents, so also the descriptions of individual documents must be structured to take account of different modes of interpersonal interaction between individual users and authors. We have to take account of variation in shared understandings, whilst recalling Lyons' comments about the present state of thinking on the problem of 'context' in linguistic study. For this reason we use the notion of 'threshold'. This, like permeability, bridges structure and operations aspects of our system.

We define threshold in the following way. It concerns the point at which the individual has adequate knowledge and experience in use of different reality systems, and consequent use of their languages, to enable him to recognise the cues which the system affords. This threshold will not be reached simply on the basis of knowledge of everyday use of the English language, though the language of some reality systems is much closer to that of everyday experience than is that of others. (Less 'linguistically' able users would derive benefit from using some segments of our system, but would not be able to exploit its full potential.) From the other 'side', the notion of threshold defines the point at which the description of individual documents provides the appropriate (contextually relevant) cues, and allows the user to draw upon familiarity with the reality systems at his command either to supply the necessary understandings or to recognise absence of rapport.

It is clear that the only acceptable approach to the description of documents for us is condensation, retaining the structure and content of ideas in the original in all important respects, together with author's terms, merely excluding least essential detail (we have formal procedures for such reduction). 'Natural' language and 'natural' structure is essential because at the level of the individual document we must intrude as little as possible between user and author, for fear of giving rise to 'misunderstanding'. This is a different situation from the level of overall structure, at which we have sought the collectively used categorisations by which, largely irrespective of the distinctive content of their individual contributions, authors have associated themselves in collectivities within the literature as a corpus. These categorisations may be expected to have familiarity through use. In addition, any given heading is defined by other headings, as well as by the set of documents associated with it. When we consider a user appraising the range of documents in a set, we may similarly expect these documents to define each other, but they will not do so in the absence of fully determined meanings. Understandings must be supplied by the user, and this can be done only by preserving the language, and hence the reality system, employed by the author.

In considering the nature and amount of supplementary contextual information to be provided (ie in operationalising the notion of threshold) a general principle in the rationale underlying the PRECIS[1] indexing system had appeal, namely that the user will test for relevance concept by concept. (In PRECIS, terms in a subject are arranged in a systematically ordered string. The string is rotated so that each term may be presented as an access point. There are special techniques to ensure that the same sense is conveyed irrespective of the rotation process. We use 'sense' here as Lyons defines the term.)

---

[1] For a general description of the system see Austin, D. Op. cit.

We are not able to use document descriptions to generate required entry points since terms individual to the document will not afford the kind of overall organisation required. If the structure of the description is to be that of the intellectual logic and categories of the document, we cannot use formally defined and ordered categories. But the notion of testing for relevance may be associated with our notion of threshold. We assume there to be a minimum amount of information which the user requires in order to classify a document in terms of the author's meaning system or to establish rapport. There will be further sets of minimum information required to make other judgments, such as relevance in terms of content, approach and so on. We may think in fact of a series of thresholds. Our descriptions, however, are not structured with any preferred order of judgments in mind. Such thresholds are simply 'embedded' in the description. This means that fairly extended descriptions are required. (Our guidelines for writing descriptions take account of alternative thresholds.)

It seems to us that, in our field, relevance judgments cannot be viewed as a simple funnelling-down process in which relevant documents are located under a given subject heading or headings, and the set of documents screened to select documents in relation to the purpose in hand (ie high/low, partial/total relevance). This oversimplifies a much more complex process. Potentially useful documents are seen to constitute the broadest set (this is the system's selection from the total universe of documents); 'actually' useful documents represent the user's first sort (including those over whose meaning system the user has mastery, and those useful in terms of substantive subject matter, along with other factors); relevant documents are those selected at a further screening stage in which the user may construct a set containing documents useful in terms of one or, more probably, a combination of these factors relative to his purpose.[1]

Comparison with other types of information system

We have outlined our system as involving multiple classification of documents into a series of alternative frameworks representing different perspectives on the subject matter of our field, rather than a single classification of our subject matter. Document descriptions reflecting the language and structure of originals allow the user to supply meanings on the basis of shared understandings with the author. They serve also as extensions of the headings, longer than the average index entry and having a formal relationship with the heading not intended with the abstract. It is considered that such a scheme will best serve users who will use the system in different ways, dependent on perspective. In principle the user is to be enabled to operate within the system as a member of a cognitive subuniverse, in the same way as he would operate upon a given body of original documents.

The principles we have outlined assume interaction on the part of user. This contrasts with input-output systems, which are based on the notion of a flow of information from source to user, via sorting and channelling processes on the part of the system, even those taking account of feedback from users.

The general idea of an interactive system is not new; it has arisen mainly in relation to computerised systems. Users access such systems directly. But it is probably mostly conceived of as a formal 'man/machine' interaction, at the 'user/system interface', if such terms can be taken as a guide to an implicit conceptual model. This contrasts with the notion of user/author or interpersonal interaction central to our model. In principle this notion should be equally valid for computer or manual systems, though our particular concern is the manual.

---

[1] A document not relevant in subject matter (eg an organisational analysis of a hospital) may be more useful for a given purpose (eg study of decision-making processes in the school) than one which is relevant in subject matter (eg on role of the headmaster).

Interpersonal interaction is sometimes equated with informal communication.[1] In this sense it is contrasted with formal systems as lacking 'logic', in contrast to the systematised nature of system-user interaction, in which the system imposes its logic upon the user. In fact, informal communication is by no means unpatterned, but the patterning may be complex, reflecting the multiple reality systems in play, unlike the single unified system of intellectual structure found in formal information systems. It has often been noted that informal communication is highly effective, but it is usually assumed that the characteristics which make it so effective cannot be employed to make formal systems more effective.[2] This is certainly the case, if the definition of a formal system excludes the interpersonal element, because insight into the complex patterning of reality systems is dependent upon awareness of shared understandings. Given shared understandings, however, a system based upon multiple structures can be used as systematically (ie is as much a formal system in our first sense) as one with a single unified structure (ie formal in our second sense). Conversely, we may say that systematic use of an information system is not dependent upon the type of control normally associated with formal systems. We prefer, therefore, to represent the options as overleaf.

---

[1] Wolek remarks that 'Interpersonal communication is generally seen to be a random and fortuitous event .... When we say that interpersonal communication is informal, we are really saying that we see no logic in it.' He feels 'a paradox exists between the necessarily formal structure of planned information systems and the apparent informal nature of interpersonal communication'. This 'may reflect only our lack of knowledge'. He feels 'it has been clear for some time that any comprehensive attempt to serve the user ... must consider the user's reliance on interpersonal communication'. (Wolek, F.W. Preparation for interpersonal communication. Journal of the American Society for Information Science, 1972, 23(1), 3-10.)

[2] Menzel notes that there are two ways in which we might go. We can either attempt to make informal communication less necessary by enabling the formal system to take over its function, or we can seek ways of channelling informal communication to better advantage. (Menzel, H. The information needs of current research. Library Quarterly, 1964, XXXIV(1), 4-19.) For the most part attention has been concentrated upon the latter (eg how we could 'use' gatekeepers). Our work falls under the former heading.

	Formal$_1$ Systematic nature of interaction	Informal$_1$ Non-systematic nature of interaction
Formal$_2$ System with single unified structure	eg Controlled language index	eg KWIC index
Informal$_2$ System with multiple independent structures	eg System such as we outline	eg Consultation of subject colleagues

The shaded cell represents the option we are aiming to provide. By constructing such a system we shall be in a position to test the assumption that it is possible to incorporate the kind of patterning which makes informal communication so effective within a formalised system, to make formal communication more effective. In so doing we avoid the circularity of designing a system to the mode of use to which a number of existing systems have aimed to condition their users.

Evaluation of system

It is an attribute of our methodology that the framework for evaluation is established by our model. Evaluation commonly involves measurement of system performance in relation to 'standard' criteria. Acceptability of performance levels is then judged either by comparison with performance of other systems, or in relation to the requirements of a given user group, expressed in terms of the same criteria. We have already questioned the assumptions implicit in such criteria. Our concern is rather with 'user performance', and acceptability of performance level will depend upon the extent to which users belonging to different cognitive subuniverses can function, without 'interference' from our system, as members of those subuniverses.

The advantage of working with the aid of a model is that we can proceed (by testing propositions about user functioning) derived from our model. To the extent that such propositions are supported, we may claim that our system is successful, that is to say it performs as intended. We are open to criticism from those who dispute the validity of our model. Such criticism would replace inappropriate conceptualisation or faulty logic in building or applying the model by something better. But, given the model, we may have confidence in our system to the extent that we can show that we have 'realised' the model. Where we fail in this respect, we must reformulate our model. The model will then indicate the ways in which we must modify our system. Every attempt is being made to make all assumptions fully explicit.

We have outlined specific hypotheses at various points in our argument. Some of these hypotheses are interdependent (eg one is the converse of another); in some instances there is entailment (a minor hypothesis will hold only if a more major hypothesis holds). All our thinking is dependent upon the assumption that individuals function as members of cognitive subuniverses. Effective evaluation will involve working out the full set of relations amongst hypotheses. At that point we shall need to conduct only a limited set of experiments to assess the validity of our ideas and of the system which reflects them. For this purpose we must test, amongst the options we have built into our system, the minimum set of hypotheses or, in other words, those hypotheses from which all others logically follow.

Such an approach is clearly economical and particularly helpful at the stage of system development and modification. It provides the basis on which to finalise a rationale which will guide eventual 'real life' implementation. At this stage we are only able to outline the method of evaluation we associate with the problem we are studying. We see this as an integral part of a methodology designed to 'break the vicious circle'. As we have indicated, we are now in process of interpreting our model in terms of our material and our users. Hypothesis construction and system construction are concurrent and interdependent processes. The relations between user functioning and knowledge structure which we realise, by the system we build, will in fact be our hypotheses about 'rules of transformation', and hence the framework of evaluation in the sense of conventional hypothesis-testing empirical research. This is what we mean by evaluation.

Practical viability

The only specialised information service in our field is Sociology of Education Abstracts (SEA). We have worked independently of this service in order to avoid being influenced, in system building, by the practical circumstances particular to SEA which guide policy and practice in SEA. It is reasonable at this stage, however, to consider the practical viability of what we propose by reference to the general way in which a service like SEA operates.

A first major area for discussion is the assumed user interaction. Whilst the fairly massive non-use of available information ser[illegible] by social scientists (sociologists of education are no exception) might seem to suggest a lack of willingness to spend time on information seeking, it seems equally reasonable to suppose that more time might be set aside if the results were felt to be more worthwhile. Our evidence suggests, as we have indicated earlier, that not infrequently material is organised in such a way that appropriate categories are lacking and there is no alternative to a cover-to-cover, volume-by-volume search. Additionally, discussion indicates that there is low confidence in and sometimes low acceptance of the way in which categories are defined and hence in the distribution of documents over categories. The result is that the conscientious user must range much more widely than the overtly appropriate categories to ensure that no material useful to him is overlooked. The less conscientious user probably does not even start. Even apart from the argument that the search process is an end in itself, the kind of situation we envisage would be likely to involve less time for the user than services such as SEA, in their present form, require. From particular points of view the outcome would be expected to be more satisfactory, and this might contribute to an overall increase in the volume of use. (We do not take the position that this is by definition 'a good thing', but in the context of this discussion of practical issues note the point merely in the spirit that possible additional subscriptions would serve to reduce costs.)

Our comments throughout have referred to academics as users. Librarians and information officers are sometimes called upon to conduct searches on behalf of the academics they serve. Our suggestions would in no way preclude this. But, since the principle of shared understandings would underlie the description and organisation of material, it would require subject background above technical expertise. Effectiveness of search will depend upon appropriate conceptual strategy rather than use of a generalised strategy (ie one considered appropriate irrespective of subject field because services are structured with a given strategy in mind). Our comments also focus upon academics rather than practising teachers. The latter would not be expected to want to derive help for their practical problems direct from research reports, even given the necessary statistical etc background to assess the bearing of findings upon their particular circumstances. Practising teachers engaging in the academic study of education (eg for a higher degree) would be expected to have the same concerns as other academics.

A second matter for discussion is the task of the indexer/abstractor and the kinds of judgments he would have to make. It will be recalled that it was envisaged that various 'thresholds' would be defined and that certain types of information would be included in descriptions. Beyond that, faced with the individual document, the indexer would be required as far as possible to produce a 'miniature' of the original. This raises the question of reproducibility.[1] Word for word replication in the case of similar items would not be expected, since we are using natural language. More important aspects of reproducibility are reliable identification of major variables, awareness and understanding of conceptual framework so that significant points are not overlooked or misinterpreted. These are intellectual judgments which would require detailed study of the document. Productivity would not be high. Such judgments too, for a reasonable degree of reproducibility across staff (as compared with internal consistency), would probably rely as heavily upon example, feel and discussion of problems as upon any formal guidelines. Subject background here too would be essential (ie a sharing of understandings with authors and users). Further judgments would be required in relating the documents to the classification scheme. Almost identical considerations and comments apply.

---

[1] Cooper has discussed the phenomenon of inconsistency (ie 'different indexers are apt to assign quite different sets of index terms ... to the same document'), and has found that 'Until a more general equation linking interindexer consistency with retrieval effectiveness has been derived, interindexer consistency cannot safely be used as a gauge of indexing quality'. (Cooper, W.S. Opinion paper: Is interindexer consistency a hobgoblin? American Documentation, 1969, 20(3), 268-278.)

Thirdly, we may consider questions of practical as contrasted with intellectual organisation. Our project is concerned specifically with the point in time bibliography, and the question of the form of supplements thus arises. The problem for an ongoing service cumulating at regular intervals is broadly the same. If we follow the logic of our arguments it would be the case that each volume, part or supplement would be organised in the light of the state of knowledge at the time of issue. Compatibility of organisation over time would be dysfunctional, given the mode of searching envisaged. One has only to study the development of thinking in the field as represented by a comparison between Floud and Halsey's[1] bibliography of 1958 and issues of Sociology of Education Abstracts since that time to see the force of the argument for reorganisation rather than cumulation. However it will be adequate if individual segments are added to or discarded from the system, as reality systems emerge or decay in the process of 'revolution'. For the rest, it will be possible to combine a reasonable degree of stability with effectiveness in use if we think in terms of a reworking of the material perhaps every five or ten years. Shared understandings will minimise the effect of increasingly bad fit with thinking in the field over the intervening period.[2]

---

[1] Floud, J. and Halsey, A.H. The sociology of education: a trend report and bibliography. Current Sociology, 1958, 7(3), 165-233.

[2] Apart from the problem of changes in thinking over time (cf Horowitz, I.L. The sociology textbook: the treatment of conflict in American sociological literature. Social Science Information, 1972, 11(1), 51-63), there is in any case a complicating factor in that there is a time lag between the development of new concepts or theories and their acceptance. By acceptance we mean legitimation in Berger and Luckmann's sense. (Berger, P.L. and Luckmann, T. Op. cit.) Kuhn argues that, in the sciences, the seal of acceptance comes from appearance of new knowledge in textbooks. (Kuhn, T.S. The structure of scientific revolutions. (International Encyclopaedia of Unified Science, Vol.2., No.2) Chicago, University of Chicago Press, 2nd ed., 1970.) Rothman has considered sociology textbooks from this point of view and finds that whilst previous 'studies verify changes in textual

continued...

A system such as we envisage is possibly open to criticism as a 'luxury' system. We have failed to find any acceptable alternative. The main call upon resources lies in subject expert time, and the main problem here is in working to guidelines, and in quality control for consensus in application of guidelines. Both Sociological Abstracts and SEA operate on the basis of participation by subject experts with editorial coordination so there is no new problem here, providing the system itself is appropriate. Cost is clearly an important consideration. Cost data being collected by the DISISS project (Design of Information Systems in the Social Sciences, based on the University of Bath under the direction of Maurice Line) will be of interest. In principle, the appeal would be to a similar market as that tapped by SEA. SEA flourishes and, as the sociology of education is more widely researched and taught, the market is expanding and the pool of subject experts on whom to call as contributors increases.

Fourthly we may consider the possibility of interchange of data with other systems. Compatibility is a key factor in this respect. If compatibility is defined to mean total and meaningful integration of documents processed in one system with those processed in another, or identical intellectual treatment of the same document in technically different systems, our conception of the relation between system and user means that our techniques will not to any great degree be compatible with those of other services. Our data set cannot be merged with another data set and retrieved by the means of retrieval assumed in the organisation of that data set, except to the extent that such organisation reflects one of the reality systems within our system and can there-

---

continued

content over time [they] suggest that there is a discrepancy between content and concerns of the discipline at the time the books are published'. (Rothman, R.A. Textbooks and the certification of knowledge. American Sociologist, 1971, 6(2), 125-127.) Crane discusses the problems for 'hard' scientists which arise from the 'disparity between "research front" classifications and those found in the literature'. (Crane, D. Op. cit.) It is the current state of thinking in a discipline that will influence users in their approaches to information systems, but there are likely to be differences in what 'the current state of thinking' means as between those in the forefront of research and others concerned largely with transmission of 'accepted' knowledge to students.

fore meaningfully accept the proportion of our documents handled by that reality system. For our part, we may similarly incorporate material processed elsewhere within a given reality system if this reality system represents a segment of our system. We must leave as an open question the extent to which independent selection and indexing will be needed to ensure that relevant material does not escape notice by either partner as a result of representation within a 'foreign' reality system.

There is no question that cooperation is desirable. It is quite conceivable that users in one social science field may find work in another meaningful and hence useful (it may be non-relevant in subject matter and yet useful). The behaviourist psychologist may well have something more meaningful to say to a positivist sociologist than does a phenomenologist sociologist. The concern of someone engaged in the academic but problem-centred study of education, who has to identify and make sense simultaneously of the perspectives reflected in psychological, sociological and other research bearing directly or indirectly on his research problem, is even greater. (Disseminating findings of research from the point of view of its bearing on practice is another question, but a related one. We have suggested that such an aim cannot be achieved in a way consistent with effective service to academics, but needs a separate and probably different kind of initiative). Current information processing techniques do not however lend themselves to the making of distinctions of this kind.

The study of switching mechanisms[1] has relevance in principle to this issue, but such work is largely being directed towards technical problems arising from characteristics of existing indexing languages. We see a much closer parallel with the cross-cultural problems of interlingual translation (information loss, misinformation or distortion, the nonsense created by overliteral translation) than with technical problems of a systems engineering nature.[1] However, little work has been addressed as yet to modes of search appropriate in cross-disciplinary contexts, and their bearing upon desirable characteristics of indexing languages.[2]

---

[1] For an account of a current project see Coates, E.J. Switching languages for indexing. Journal of Documentation, 1970, 26(2), 102-110.

Finally we deal with a question of future policy. On the basis of Brookes' analysis of the UNISIST report, discussed earlier, it would seem that integration could well lead to uniformity of a mindless sort. We can only echo the reservations which Brookes attributes to social scientists. Administratively viewed,[3] the advantages may be considerable. But we would argue that intellectual considerations should determine the modus operandi of such a scheme, or of similar but more localised schemes such as EUDISED (European Documentation and Information System for Education), which is nearest to our interests. We believe that there may be overriding reasons for preserving variety (and even for adding to it) in the intellectual schema underlying social science information services. It is possible that this variety represents not only individual rationales[1] but also conceals regularities of ideas which are not immediately apparent, because in considering them in relation to preconceived notions or values (finding them perhaps idiosyncratic) other patterns are masked.

---

[1] The problem of the merging of data bases highlights this issue. Green has studied 'Some problems arising in the combination of discipline-oriented systems [DIR] for the purpose of constructing a specialised retrieval system [SIR] for a multidisciplinary topic'. He found it was not possible 'to produce a suitable indexed data-base ... by the simple process of conjoining the bases of a set of discipline-oriented systems', though 'it would still appear possible to use many of the facilities provided by the larger systems, such as document accession, machine manipulation, and clerical organisation'. His investigation demonstrated that 'The classification and indexing procedure of a DIR system is crucially influenced by its orientation'. He notes that 'Since for multi-disciplinary topics the subject of interest may be secondary to the main field of the document the indexing employed by a SIR system must also be biased'. The special area with which he was concerned was that of Numerical Methods. (Green, C.D. Some problems in the indexing of specialist material drawn from several disciplinary systems. Journal of Documentation, 1972, 28(1), 37-43.)

[2] In more general terms, Lynch notes that 'It is surprising that so little attention has been paid thus far to methods of index use ...'. (Lynch, M. Computer-organised display of subject information. Indexer, 1971, 7(3), 94-100.)

[3] We have not dealt with issues of this kind in this paper, but we believe that sociological study of formal organisations would provide useful insights into the strengths and weaknesses of alternative forms of supporting bureaucratic organisation for information systems.

We can best explain this point by describing a study[2] which investigated the consistently chaotic state of medical reports noted by a group of researchers wanting merely to extract certain simple categories of personal data on patients. Their purpose was to study criteria of selection for entry to a psychiatric clinic. It was found that the records were kept on a 'contractual' basis (ie the social contract of the clinic with regard to proper care of the patient) rather than an 'actuarial' basis. The major, though not the only, concern in record-keeping was that informed readers should be able to reconstruct a case history from the point of view of 'proper care'. Omission, overlap or duplication of records, contrary to the case in actuarial recording of data, often held for a reader who was familiar with the routine working of the clinic a significance of a kind which was professionally relevant even though superficially the records might seem ill-kept.

---

[1] At the very least, the unanimity with which social scientists stress the importance, for their work, of a diversity of intellectual approaches must give us pause for thought. In a general context, for instance, the editors of the *International Encyclopaedia of the Social Sciences* (17 vols. New York, Macmillan-Free Press, 1968) indicate that they 'considered whether it was possible to develop a conceptual outline of all the social sciences', but 'it became apparent that this was not possible'. They comment on the 'many analytical levels' to be handled. Reviewing this work, Smelser approves the decision to depart in practice from a purely 'analytic emphasis', and to present the world also in 'fairly concrete categories'. He approves also the recognition, by the inclusion of biographies, of the fact that 'knowledge is to some degree associated with individual persons, "approaches" and "schools"', such that 'some knowledge ... is important because someone said it in a certain way rather than because it has met the test of scientific verification'. (Smelser, N.J. Review of the International Encyclopaedia of the Social Sciences. *Social Forces*, 1970, 48(4), 534-537.) Similarly, Murdock et al., in describing the fourth edition of the outline of cultural materials used by the Human Relations Area Files in New Haven, stress that 'in earlier editions they attempted in several instances to institute classificatory innovation that seemed sounder on theoretical grounds than categories in general use', but 'they were compelled to abandon the innovation because the attempt to press the data into a new mold invariably necessitated splitting up passages in the sources and distributing the parts so widely that the context for each individual item evaporated'. (Murdock, G. et al. *Outline of cultural materials*. (Behavior Science Outlines, Vol. 1.) New Haven, Human Relations Area Files, 4th rev. ed., ?1966.) In a more specific field, for instance, the Committee for the *Index of Economic Journals* describe their attempts to take account of the 'structure of economic thinking' in discussing their classification schedule. There seems to be a consensus in the stress laid upon the *functional* nature of diversity within and across fields (cf the references to 'schools', to 'context') which is more than coincidence, and which suggests that the information loss associated with standardisation may be too high a price to pay for what is gained. By restricting choice we reduce use of an information tool almost to finding for finding's sake.

[2] Garfinkel, H. 'Good' organizational reasons for 'bad' clinic records, in Garfinkel, H. *Studies in ethnomethodology*. Englewood Cliffs, New Jersey, Prentice-Hall, 1967.

When these researchers attempted to obtain and translate the (for them) appropriate data in terms of the requirements of their pro formas, they were felt to give a false picture, even when researchers were allowed to make inferences. (Most researchers are familiar with this as a problem in questionnaire studies, when it becomes evident that one has been either asking the wrong questions or asking the right questions in the wrong way.) The reason was that the structure that was being imposed upon events was derived from preconceived assumptions rather than from the events. There were in fact good organisational reasons (in terms of patient care) for so-called bad records. Once this was understood, their systematic nature became evident as the records were studied again from this point of view. Good patient care was clearly of prior importance, quite apart from the very marginal utility of actuarially kept records in relation to this end, though this did not exclude the keeping of as much actuarial data as was necessary for good patient care.

Just as doctors have contractual obligations to their patients, so users have similar obligations to the field in which they work. Just as the nature of the contract in the situation studied by Garfinkel determined the relationship between contractual and actuarial record keeping, so discipline considerations may properly guide policy with regard to the management of knowledge on behalf of users.[1] The working out of such a 'contract' or definition of 'good service', in consultation with subject experts, has guided policy thinking in our project.

There was a strong indication, in the early stages of the sociology of education project, that users in our field thought rather differently from ourselves in this respect. Their terms of reference appeared to be different. For this reason we have developed a methodology to take account of differences in thinking. In order to contrast the nature of users' assumptions with our own, and to reach agreement on the assumptions which should determine appropriate information processing techniques, we have been forced not only to investigate their ideas but to make explicit and to justify our own. Our conclusions about the validity of our ideas with regard to the principles which should guide information processing have led us to a redefinition of a number of notions central to our thinking. User approval of our eventual products is the ultimate test. Success in these terms will in the first instance depend upon evaluation testing the logic of the reasoning guiding us. This paper has set out the arguments which have forced us to revise our ideas, and we have traced and illustrated the implications for research methodology.

---

[1] In the linguistics field, a restricted circulation report on the LINCS project stresses that the project has placed a special emphasis on a mandate from users. 'Its mandate to pursue the objectives embodied in the LINCS project has arisen through CAL's continuing relationships with the major professional organizations in the language sciences ....' (Roberts, A.H. The system of communication in the language sciences: present and future, in Nelson, C.E. and Pollock, D.K. eds. *Communication among scientists and engineers.* Lexington, Mass., Heath, 1970. (Final version of LINCS Report # 19-69P.))

III 3. (ii) Sociology and information science

Paper prepared by Working Party of subject experts
(see p.9): draft as at Easter 1973.

Abstract

The growth of the social sciences has necessitated the development of formal systems of information storage, retrieval and dissemination over and above any informal and personal contact between individuals. The contribution of information scientists, in their attempts to structure social science information, raises certain questions about the epistemology of the social sciences. The writers examine these and some of the major problems related to the communication of information in sociology. They postulate a model derived from recent thinking in the sociology of knowledge, ie that the construction and validation of knowledge is itself a social process. Some possible implications of this approach are investigated.

Introduction

The form of communication of information within a subject or discipline develops according to a more or less standard pattern. Initially, the learned, critical and curious in a field are few in number, and, if widely dispersed, they keep in touch by letter, semi-formal gatherings, or occasional publications. As their numbers grow they form institutions for the promotion and publication of their specialism, and a less personal and more formal system of communication begins to develop. Subsequently, specialised journals are established and, later still, secondary guides to the literature such as indexes and abstracting services appear. It must be stressed, however, that these later developments do not supplant personal communication, and the relative importance of formal and informal systems may vary.[1] The formal systems may be regarded as supplementing the informal when the latter prove inadequate. In what follows we will explore the implications of the development of formal systems for the diffusion of knowledge in the field of sociology in the belief that they have relevance, not merely for sociology, but for the social sciences in general.

Sociological literature is accumulating at a vast rate. Pease and Rytina report a doubling of the number of American sociology journals in the fifteen years up to 1968.[2] Some people have argued that we face a crisis from the sheer mass of literature and that consequently there must be some urgency in the development of information retrieval systems. Against this, Bar-Hillel[3] argues that increasing specialisation will narrow the researcher's field of concern so that the amount of literature that the individual needs to master will not grow. It may be a point in support of this view that sociologists do not make extensive use of those available information retrieval systems that are designed for them. Neither of the above positions, however, completely describes the situation.[4] To create an information retrieval system does not necessarily help to solve the information problem. On the other hand, specialisation does not mean that formal services are not required. Most published material is not organised in a manner which makes it possible for the specialist to locate relevant items immediately without selecting from more general work. Having identified his 'special' material a further selection process is required to eliminate trivia. Moreover, even if media were organised in such a way as to promote ease of location for the specialist, it does not preclude the possibility of useful material being located in non-specialised sources.

We are assuming, therefore, that whatever the investigation in hand, and whether or not there is a perceived need for information, there will be information in the literature which will be helpful in some way.[5] The problem is to ensure that it can be located easily and accurately.

If we assume that the total bulk of material to be searched is too large for a given user to look at all of it, then any retrieval system will have to structure it in a way that allows him to identify relevant subsets, through relevant entry-points. But the very existence of structure places limits on the accessibility of information. A given document may have relevance for a particular user in many different ways. Its theme, its methodology, its theoretical assumptions, and the terms it uses are all characteristics that a user may or may not find relevant. In the process of structuring documents in a retrieval system, only a limited number of these characteristics can be used to place a document and thus only those characteristics can be used in searching. Not only are there many ways in which a document might have relevance for a particular user, but also the very grounds on which documents might be considered relevant can shift in the course of the search.

So the creator of a document structure has a double task. He has to try to maximise the availability of the various ways in which a document might have relevance for a user, but at the same time has to produce a description of that structure in such a way that it can be understood by any user at any time. This description is the gateway for the user to enter the material. In effect, he is faced with an impossible dilemma. The individual user would, ideally, like to be guided through the total bulk of the material along pathways defined by his own system of relevance and his own mode of working. On the other hand, the sheer number of such users implies a system, common to all users, which can be codified as a set of instructions for all to follow.

Such a system can be a powerful influence on the communication between those who write documents and those who read them. The reader may only encounter a writer's work through an intermediary (the creator of the information retrieval system) having decided that certain aspects of the writer's work are the 'key' aspects. Moreover, the grounds on which key aspects are chosen may not be apparent to the user, however carefully the system is described. There is always a tacit element in any description, and hence there is always the possibility of incomplete understanding. Furthermore once a system is in use, there is little possibility of direct contact between users and the creator of the system, and any doubts or problems that the former may have may produce nothing more positive than an individualised and unarticulated feeling of dissatisfaction. This puts the intermediary, whether he likes it or not, in a position of considerable power. The direction of theoretical development and of research can be influenced profoundly by the imposition of a certain structure on the documents that provide a major channel of communication between members of a discipline. The process of communication may become distorted or restricted when those who are defined as appropriate persons to create and control information retrieval systems are not themselves actively engaged in the intellectual development of the ideas they are attempting to structure.

By distortion we mean that the meanings imputed to documents take little account of the standpoint of the discipline within which both author and user are working. This will result in the creation of sets of documents containing a high proportion of material which experts in the discipline will perceive as inter-related in only a trivial sense, if at all. For instance, anthropological, behavioural-psychological and sociological studies of cultural deprivation would not usefully, for the members of any discipline, be brought together as a single set. This is perhaps an obvious example. The meanings of other terms may be less readily recognised by those not actually engaged in the use and development of the concepts of a given discipline or disciplines. For instance, documents presenting social-psychological studies of occupational careers, where the unit of study is the individual person, would not in general be sought by sociologists engaged in the investigation of the career patterns of occupational groups. This latter type of study is further distinguished from studies emphasising status sequencing and extending the notion of career beyond the context of occupations to, for example, the career or life cycle of the mental patient as in Goffman's study Asylums (Chicago, Aldine Publishing Company, 1961). Other sociologists have extended the notion to interpret the life cycle of, say, the family or the organisation as a 'career'. The danger is that, even if these different meanings of 'career' are kept distinct and clearly defined in a classification or indexing scheme, non-specialist indexers may not be able to impute to the term the particular meaning it has in a given document. When this is the case, related material will be separated, and may become distributed under a number of headings in ways which the specialist will be unable to predict.

Cont'd over....

By restriction we mean that sets of documents will not be constructed in such a way as to bring together material which, though possibly not dealing with the same substantive problem, is nevertheless conceptually related. For instance, documents describing empirical work in the context of, say, a hospital and a family might not be brought together, despite the fact that the authors of both documents looked at interpersonal relationships in their specific context using an implicit power model; thus the user might well be cut off from documents which could be usefully translated into his frame of reference.

Another aspect of imposed structure is that relationships amongst sets of documents are indicated by cross-references in a way which channels the user along predetermined paths through the total corpus of documents, thereby making his ideas shift in one direction rather than another. This is, of course, true of any system, but it merely illustrates another way in which the imposition of structure can affect, albeit in unknowable ways, the direction of theoretical development and research.

In saying this, we are not challenging the notion that there exist specialised skills in the creation of information retrieval systems, but we do challenge the notion that these skills can be segregated and exercised independently of skills in the subject field they serve. Indeed we are equally suspicious of leaving the structuring of documents to any closed group of 'experts' from within the discipline. The development of information services should be an active and on-going partnership between those who possess the skills of information science and those who understand the discipline within which the documents are written.

Responses to the information problem from within the social sciences and by information science

Subject specialists make their information available through a wide range of books and journals. However, these have long been recognised as too diffuse in their aims and interests to enable efficient communication within growing and diversifying fields, and do not go far enough to ensure ease of access as well as availability of information about current work.[6] Abstracting services, which are essentially non-selective and non-evaluative, are complemented by critical reviews and overviews of research, or 'state of the art' reports such as one finds in Current Sociology. The latter are concerned less with the completeness of the record or the presentation of substantive findings than with the identification of significant pieces of work, with significant trends or with the analysis of perspectives. A large number of sociologists appear to be content with the present situation and with the part played by sociologists in evolving information services, particularly when this is coupled with the informal channels of communication which everyone uses.

Information scientists view the situation somewhat differently. They would claim that a documentation system which has developed by means of a number of separate initiatives suffers from duplication of coverage, from gaps in coverage, and from a confusing variety of methods for organisation and retrieval which the user has to master.[7] By and large it is because the growth of informal communication has not kept pace with the increasing numbers of specialists in any given field that a new profession of information science has come into being, particularly in science and technology. Recently the information scientists have turned their attention from the physical sciences to the social sciences, in particular to psychology.[8]

What is overlooked, in our view, is that information control should represent part of the organic growth of the discipline of sociology. The initiation of formal systems represents a stage in the institutionalisation of communication processes within the discipline, and their increasing differentiation of function is a reflection of perceptions of needs by those in it. At the same time such systems undoubtedly exercise influence upon the discipline itself. An obvious example is the power of journals in setting and maintaining academic standards, conferring prestige upon individuals and, through their selection policies, controlling the availability of new ideas.[9] Crane,[10] for instance, suggests that an important factor in the selection of articles for journals is not so much personal ties between editors and authors but similarity of academic training, methodology, theoretical orientation and mode of expression.

Less obvious is the power of abstracts and indexes to influence thinking. In a review of information needs and uses, Crane[11] hints at some of these problems. She highlights research which shows that classification schemes used by indexing and abstracting services tend to lag behind those currently used by scientists and that problems of terminology are magnified when a scientist seeks work in a new field or a field other than his own. In 'translating' his problems into the terms used by the information service he distorts the meaning of his ideas. Thus the documents he retrieves may be a poor match with his requirements. Additionally he may fail to locate relevant documents because the categories which describe his problem do not exist within the scheme of organisation used by the information service. Crane goes on to state that '"the relatively mediocre results" of automatic text analysis procedures appear to be a consequence of the variety of ways in which terms can be used in scientific documents. It seems likely that further progress in the development of devices of all sorts to speed information-seeking will come about only as a result of basic research on how scientists use ideas, which is in turn related to understanding how knowledge develops in an area'. Such problems are particularly marked in the field of sociology where meanings of terms change not only as sociology develops but as society itself changes. Problems of terminology are apparent not only when an individual attempts to cross discipline lines but also when he wishes to locate his problem via the terms of another theoretical viewpoint within his own discipline.

Services organised by specialists for similar specialists, implicitly if not explicitly, are sensitive to such considerations and are subject to appropriate controls in terms of accountability to colleagues within the discipline. To some extent, control over the diffusion of knowledge is now passing from social scientists themselves to professional information scientists whose academic background may not be that of the field they are serving yet who will hold a position of considerable power within the field.

The occasion of the present writers' awareness of the problems of information processing was an attempt to develop a system for compiling and constructing an index to a bibliography of material relevant to the sociological analysis of education. Consideration of the principles upon which current systems of indexing are based led us to question some assumptions which appear to underlie information scientists' models. We started to look at the epistemological basis of sociology itself, to see why we had such difficulty in indexing sociological material, and in doing this we realised that information storage and retrieval was itself a subject of sociological enquiry. We realised that what sociologists had to say about the social construction of knowledge had implications for the general construction and codification of all knowledge.

With regard to these implications, we have looked at some theories and procedures of information scientists to try to discover areas of convergence and divergence. It is impossible in a brief statement to do full justice to the diversity of approaches currently being employed by information scientists. The models to which we refer in this discussion are those of two influential information scientists who have made explicit their beliefs concerning the construction of knowledge and the implications for the design of information systems. We are not claiming that these models are representative of all information science but feel that the assumptions on which they rest underlie a number of attempts to produce information systems.

Austin[12] has said that there are at least four levels involved in the creation and transfer of information:

'(a) the phenomena observed;

(b) the images in the mind of the observer;

(c) the expression of these in writing (and other media);

(d) the extraction and coding of elements in these media to function as retrieval 'keys'.

'As documentalists we are mainly concerned with only the last two of these, and although concepts such as reality and cognition are clearly of interest we are not concerned with matters at those levels except insofar as they might affect the performance of the system we devise.'

We agree with this analysis, but feel that in practice the documentalist works as if the first two levels, in fact, do not affect the performance of systems. We, on the other hand, feel that the first two levels do significantly affect the performance of systems and that what sociologists require at Austin's fourth level of information processing is to know how the author of the document perceives the relation between the first three levels.

Foskett[13] also considers the relationship between these levels. He has said, 'I agree that the cognitive behaviour of individuals is "frozen" in documents but it does not acquire meaning only as other human beings impose meaning on it. It surely has its own meaning, which is true in so far as it approximates to what is actually taking place in external reality. Knowledge is certainly created by human beings, but it is on the basis of what is "out there", and I am totally at a loss to understand how the notion that a structure of ideas exists in a corpus of documents is misleading. If writers have any purpose at all, then it is surely to embody a structure of ideas and ordering of knowledge in what they write.'

The following assumptions may be seen as underlying the statements of both these writers:

1. Knowledge is structured, and possesses criteria of meaningfulness which are independent of the knower.
2. There is a referential relationship, which is constant and direct, between the ideas which comprise knowledge and the linguistic constructs by which these ideas are expressed.
3. The structure of ideas in single documents can be fitted into a common framework that covers any corpus of documents.

We would criticise these assumptions as being too limited, both as regards sociology in particular, and the social sciences in general. In the next section we explain in more detail the limitations imposed by these assumptions and illustrate these limitations with the problems we have encountered in our attempts to index sociological documents. We do this in the hope that it will initiate further discussion between sociologists and information scientists, so as to allow us to devise a joint solution of the information retrieval problems in sociology.

A sociological approach to information

Many views of knowledge have been held at one time or another by sociologists. One major tradition of possibly decreasing importance is instanced here: that of classical scientific method. This method, with its inherent assumptions of causality, 'natural order', and quantification has tended to operate a policy of exclusion, resulting in the view that approaches which are not based on this positivistic method are inadequate. The collection of valid data depended upon detached observation according to formal rules of procedure. This could then be used to refute or give plausibility to hypotheses deduced from theoretical schemes which themselves were thereby put to the test. This methodology was believed to be shared by all scientists. Early anthropologists and sociologists based their work on the assumption that the relation between terms and their referents was somehow absolute. For example, to view poverty as a matter of personal perception on the part of the poor rather than as something related to an absolute norm, the poverty line, would have seemed incongruous to the early social reformers. This reliance on a conception of meaning which assumed that it was a property of objects or events was a result of the transfer to the social sciences of a viewpoint which dominated the physical sciences.

This conception of meaning as independent of the knower, we have suggested, appears to underlie at least some of the current approaches to information handling.

In sociology there has never been a total commitment to such a notion of meaning as a sufficient basis for the interpretation of social phenomena. Weber, for instance, found it necessary to distinguish between subjective and objective states of human behaviour. It is only in the ascription of individual motives and assumed goals that the observer can come to understand behaviour in such a way as to make possible any later discussion of the causes.[14]

Recent writers in the sociology of knowledge[15] have moved away from the notion that the observable and non-observable aspects of human action can be considered separately and have meaning in themselves. They see the subjective and objective dimensions of experience as simply different ways of viewing the same world, and both are essential for an understanding of any situation. This approach we believe to be best for an understanding of the information problem faced by sociologists.

In essence, what starts as subjective experience comes to be viewed as objective knowledge through communication and subsequent consensus. The researcher has to transform his experience into some form of symbolic representation that others will understand. That is to say, he has to use, or modify, an already existing and publicly available framework of ideas and their associated symbols. However, the relationship between ideas and symbols is never completely explicit, but is maintained by direct communication amongst individuals who, in groups, develop their own pragmatic rules for agreeing on what is, and what is not, valid knowledge.

Different theoretical perspectives, then, are maintained implicitly by different sociological communities whose members have learnt a particular way of looking at the world. New knowledge, which starts as personal experience, becomes objectified by framing it in a particular perspective and validating it according to the rules of that perspective. For members of that community the knowledge is then a 'description of reality'.

We have come to realise that the subsequent framing of documents by an information system is a further process in the objectification of knowledge. Members of a community that use an information system are thereby acquiring a specialised 'description of reality'. What they see in the documents will be, as we have stressed earlier, a product of the system used. For example, when two documents are put into the same category because the same term occurs in both, then users of the information system have to assume, along with the creators of the system, that the documents are dealing with some common entity to which the word refers (see assumption 2, page 10). But, as most sociologists would agree, this is to oversimplify the nature of many, if not most, terms in sociology. The meanings of terms in a document are mediated by the theoretical perspective in which they are embedded. Understanding of the perspectives in a discipline is acquired through communication and the development of shared understandings, which are, in part at least, implicit. That is why we are particularly concerned that the creation and use of an information system should be an active partnership between the practitioners of a discipline and information scientists. Moreover perspectives change with the development of thinking in a discipline, so that this partnership should be an on-going one, and not merely a matter of consultation at the initial system design stage.

We suspect that the preceding argument applies to all disciplines, not merely to sociology, but there is a further problem specific to social science fields with the assumption that terms in documents can be defined solely by their referents. Social science concepts cannot be seen as corresponding to a single referent because they cannot be dissociated from their everyday meanings and the everyday life which forms the subject matter of the social scientist. The researcher is a part of his subject matter and consequently the knowledge that he creates owes something to the nature of his involvement in the society he is studying.

Those who seek sociological literature may therefore wish to locate themselves and their work in terms of a community of others who hold the same, or a different, view of the phenomenon in question. Such a distinction might be the only criterion of relevance that a user will employ, but more probably it will allow him to extract a set of documents on which further criteria will be employed to decide relevance.

The user is trying to perform two basic tasks when approaching an information retrieval system. Firstly, he is trying to locate his initial conceptualisation of a problem in an existing body of knowledge. That is, he is trying to identify others who have talked about his kind of problem in a way that he understands and with which he sympathises. He is trying to establish a genealogy for his ideas. Secondly, he is looking for alternative approaches to the same problem, so that he may take the work of others and translate it into his frame of reference. This may involve little more than taking the data of others and reformulating the inferences to be drawn from the data.

If the views that particular groups of sociologists have of the phenomena they are studying can be important in the literature searching process, it seems reasonable to assume that any index to sociological literature ought to take such differences of view into account. An index cannot reflect all possible perspectives, in that, in the extreme case, the individual's view of social phenomena is unique. But, on the other hand, it should aim to reflect all major differences amongst groups of sociologists. Since documents acquire a relevance by virtue of their position within the corpus that is the retrieval system, this corpus should take account of alternative perspectives in its structure. Each user has to create his own structure from the materials available. Nevertheless, at any given point in time, it may be possible to give rough boundaries to theoretical perspectives and so provide a structure to an index using certain basic dimensions likely to be recognised by the majority of users.

The fact that what is seen as a perspective by a sociologist cannot be made fully explicit, even though he feels he can recognise it when he meets it in a document, creates problems in putting boundaries to perspectives. However, we believe it may be possible for a user to give a provisional 'location' to his perspective by considering where he stands in relation to various basic assumptions.

Perspectives vary in their assumptions about such things as the nature of man, the nature of society, the topic being studied, the methods of study, and even the epistemology involved. Any or all of these may be important. Together, they illustrate our claim that what sociologists need at Austin's fourth level of information processing is to know how the author of the document perceives the relation between the first three levels. (See page 192)

A given perspective may well share many of its assumptions with other perspectives so that they overlap. We feel that the best hope is to attempt to operationalise the basic assumptions held by different communities of sociologists in terms of positions on different dimensions. The perspective of each community could then be represented by a discrete set of such positions which could be used by the individual to locate himself and his problem.

By and large, the indexing systems in use at present do not help the user to locate a document in terms of a particular theoretical perspective. They rely on key words or descriptors to provide the elements by which a document is indexed, and our argument is that the meaning of such words is only given by seeing them in the context of the whole document, and by seeing that document as part of a corpus of documents produced by a particular community of sociologists. This community, let it be added, will have imprecise and shifting boundaries, and its theoretical perspective may never be completely explicit.

If different groups of sociologists use common words, but mean different things by them, it might seem possible to tag each term by a set of procedures for identifying the perspective to which it belongs. To some extent, this may provide a partial solution. For example, if we are faced with the term 'social class' we can start to locate this term by asking whether, in addition to referring to a group of people, it also implies a corporate sense of identity for that group which opposes it to other groups. A sociologist wishing to find documents using a Marxist approach to social stratification would only take those documents for which this was so.

However many questions of this sort we might ask, we would never get a complete identification of 'Marxist' documents, not least because the boundaries of Marxism are indefinite. Nevertheless, by reading the whole document, a trained sociologist would be able to make some sort of judgement. Even then the judgement would not be an absolute one, eg "This man is a Marxist", but more of the form, "This man is talking about class in a way which allows me, more or less, to compare my ideas with his". The distinctions between the possible meanings of class are not then distinctions that can be made at the level of empirical referent or linguistic rules. They require an understanding of how each author views the world and the processes by which he sees society developing. Such distinctions are only really evident from the total output of authors. However, clues to their views of the world can be gained by looking at the conceptual relationships between terms they use to describe their world, the connections they see between their work and the work of others, and to a smaller extent the particular problems they choose to investigate.

Conclusions

Throughout this paper we have sought to emphasise and explain certain salient points. These are here summarised not as universals, but in relation to those specific issues under discussion:

1. We disagree with the thinking currently guiding the development of retrieval systems on the grounds that knowledge is thereby handled as if it were independent of the knower.
2. Knowledge is largely dependent on the social context in which it is generated and used.
3. Perspectives comprise, and develop in terms of, shared understandings which are often in part implied and not accessible to an external observer.

In addition we would stress that the results of some three year's investigation and contemplation of the problems of information storage and retrieval lead us unequivocally to the conclusion that really effective retrieval systems depend on an active and on-going partnership between information scientists and subject experts. We regard this paper as an opening contribution toward the establishment of that relationship.

Notes

1. A particular strength of informal communication is that it draws upon personal knowledge and depth of insight into the minds and work of individuals in a community and allows more immediate and accessible interaction. As Menzel states (Menzel, H. Planned and unplanned scientific communication, in Barber, Bernard and Hirsh, Walter eds. The sociology of science, Toronto, Collier-Macmillan, 1962, 417-441), it provides the service of 'screening, evaluating and synthesising' material. The existence of some highly organised structures for informal communication is well-documented in the natural sciences. Price (Solla Price, D.J. de, Little science, big science, New York, Columbia University Press, 1963, pp. 62-91) uses the notion of 'invisible college' to describe groups of elite scientists who keep in regular contact with each other through conferences and personal communication. Such structures might be far more efficient than any formal system but nevertheless are limited by the sheer number of people it is possible to keep in contact with. A more serious limitation is that accessibility to informal communication networks is differentially distributed and thus markedly inefficient for some workers. In the United States National Research Council report on the communication system in the behavioural sciences, (Communication systems and resources in the behavioral sciences, Washington, D.C., National Academy of Sciences, 1967) it is noted that the existing communication system works well for some scientists, most of the time, but is inadequate for others. The scientists that find the existing systems to be adequate are more likely to be older and more established, to have many professional contacts and to sit on committees and governing councils of professional bodies . It would seem that as

far as possible sociologists still use the informal system in preference to the formal. (See Line, Maurice et al. Investigation into information requirements of the social sciences: research report No.1, Bath, Bath University Library, 1971.)

2. Pease, J. and Rytina, J. Sociology journals, American Sociologist, 1968, 3(1), 41-45.

3. Bar-Hillel, Y. Is information retrieval approaching a crisis? American Documentation, 1963, 14(2), 95-98.

4. Holt, Charles C.C. and Schrank, William E. Growth of the professional literature in economics and other fields, and some implications, American Documentation, 1968, 19(1), 18-26.

5. Irrespective of the objective situation, in terms of sheer volume of literature, sociologists need not necessarily feel themselves to have an information problem which they must come to terms with if their work is to proceed. Unlike the situation in the natural sciences, it seems unlikely that sociologists face a real problem of duplication of work even if they fail to keep up with the literature. This fact arises because of the wide variety of 'accepted' methods and theories which can be brought to bear on any given problem at any point in time. However, efficiency in the sense of non-duplication is by no means the only basis on which information need can be estimated. Some sociologists, no doubt, require information for a variety of purposes at different points in time.

6. It is ironical that the journal now contributes to the information problem whereas originally it was intended to overcome it by providing a shorthand version of more lengthy, often previously published work. The growth of sociological journals has been enormous, and Ulrich's International Periodicals Directory (Ulrich, C. F. Ulrich's International Periodicals Directory: a classified guide to a selected list of

current periodicals, foreign and domestic. New York, London, Bowker, 12th edition 1967-8 and 3rd supplement 1969) lists some 371 journals in its general sociology section. The journals themselves represent chiefly the activities of learned associations such as the British Sociological Association (founded 1951) and the American Sociological Association (dating from 1905), university publishing houses and private interest groups. These publications have long been recognised as being too diffuse in their aims and interests alone to enable sufficient communication within the growing and diversifying field, and in particular much of the work remains embedded in the journals. The major abstracting service in the field of sociology is Dr Leo Chall's Sociological Abstracts which began in 1952. A British venture in a limited field is represented in Sociology of Education Abstracts begun in 1965. Even in a short period of time and in the limited area, this latter is again proving inadequate from the 'ease of access' point of view for the needs of its users.

7. Line, Maurice. Op. cit.

8. A series of detailed studies for information exchange has led to the design of a national information system for psychologists (NISP) in the United States. A brief review of the research leading to the proposals for the national system is provided in a paper by van Cott (van Cott, Harold P. National information system for psychology: a proposed solution for a pressing problem, American Psychologist, 1970, 25(5), i-xx). The proposals which involve the circulation of unrefereed materials to A.P.A. members have been the subject of much controversy (Boffey, Philip M., Psychology: apprehension over a new communication system, Science, 1970, 167, 1228-1230). The system's supporters see it as an imaginative effort to cope with problems arising from the rapid growth of scientific literature which will help to advance psychology as a science and as a profession. Its opponents (see for example

Loevinger, Jane NISP: noisy signal in psychology, Transaction, 1970, 7(7), p.10) see it rather as resulting in a fall in the amount of useful information actually transmitted and eventually to a decline in the quality of research being done. Its proponents see a system developed by users for users but the critics can only see the language of marketing, product packaging and cost per unit readership. As yet it is too early to assess the value of this service in terms of perceived usefulness and actual use made of it by psychologists. We may note in addition that in this country, for instance, a research team is addressing itself to the design of services for the social sciences in general (Design of Information Systems in the Social Sciences, directed by Maurice B. Line, Librarian, National Central Library). A European network of services to education is also actively being planned. A national system serving educationalists in the United States has been in operation for some years.

9. Just as the editors of the journals of sectarian movements (eg Catholic Apostolic Church, Society of the New Church, Christadelphians) control, to some extent, the ideology of the movement, so the editorial boards of academic journals exercise control over the content of the journals and the subsequent direction and availability of new ideas.

10. Crane, D. The gatekeepers of science: some factors affecting the selection of articles for scientific journals, American Sociologist, 1967, 4(2), 195-201. See also, for example, the debate in Journal of Health and Social Behaviour, 1970, 2(4), 327-329, in the form of two letters. In the first Edward Sagarin attacks Howard S. Becker's review of David Matza's Becoming deviant on the grounds that the book is, in fact, a homage to the reviewer. Becker's reply is a refutation of this. Whether the argument is valid or not Sagarin makes the interesting point that an editor of a journal can do one of four things

with a book: a) ignore it; b) give it to review to someone unqualified and inexpert; c) give it to someone precommitted in its favour; d) give it to someone precommitted against it.

11. Crane D. Information needs and uses, in Cuadra, C.A. ed. Annual review of information science and technology, vol. 6. Chicago, Encyclopaedia Britannica, 1968, 3-39.

12. Austin, D.W. Personal communication.

13. Foskett, D.J. Personal communication.
Foskett has commented on his particular position, after reading this paper in draft: '... what I have said is that knowledge is structured, and that its criteria of meaningfulness are derived from the structure of the universe, and not from the structure in the mind of any particular individual. If this were not so, we could never arrive at any kind of consensus which, manifestly, we are doing all the time.'

14. For example, see Schutz, Alfred Concept and theory formation in the social sciences, in Emmett, Dorothy and McIntyre, Alisdair eds. Sociological theory and philosophical analysis, New York, Macmillan, 1970, p.3.

15. For example, see the work of Berger and Luckman; Schutz; and Holzner.

III 3. (iii) Responses to our thinking by subject experts and information scientists: correspondence June-December 1972

This paper reports on general points emerging from correspondence stimulated by the circulation, for comment, of our paper Methodological issues and information science.

(1) Responses by subject experts

We were greatly encouraged on the whole by comments received from subject experts. The general tone may be summarised in the words of one correspondent as indicating a view of the work as 'valid, important and exciting'. Since our correspondents were known to be working in different academic traditions, this was all the more satisfactory.

We had sought to take account of the wide variety of schools of thought, and this suggested that we had been generally successful. We had, however, anticipated that our position might not be acceptable to some proponents of two particular schools of thought. We had in mind the thorough-going positivist and the thorough-going phenomenologist. The first would be expected to hold the objectivist position we challenge, not merely as a working basis for research, but as a matter of ontological belief, and hence would reject our position. The second would be likely to reject our position on the grounds that, to realise our ideas, it would be necessary to adopt some working objectivist assumptions; this would be a contradiction of our ideas.

Thorough-going positivists are relatively rare, and we have no responses based on this position. We are able to report a response reflecting a phenomenological standpoint. From this it would seem that we must accept that there is a small group of users whom we cannot serve, for the reason that their standpoint is such as to deny the validity of even the notion of an indexing system:

'(a) A feature of the construction and use of an indexing system is that its "good sense" is a practical accomplishment, depending upon the ad hocing practices employed by relevant members. Since these ad hocing practices will always assume priority over any set of coding rules (indeed will be used to display that any categorisation is in-accord-with-a-rule), the test of acceptability to professionals, which you mention, is the only test for an indexing system.

'(b) It follows from (a) that, in constructing an indexing system, a researcher makes a choice about how he will attend to the social world. That is to say, that the auspices which provide for the possibility of doing indexing produce a preconstituted world awaiting explication by means of rule, i.e. the "external object world/normative" paradigm of your own classificatory scheme. Thus, despite a rejection of an objectivist theory of knowledge (and of positivism?), the task depends for its success on very similar auspices.

...

'As the tenor of my remarks will indicate, my view is that indexing is a practical activity done by and for competent members (i.e. persons able to see the good sense of a particular classification by invoking the appropriate ad hocing practice). I suspect, then, that sociologists with my particular concerns will find most relevance in the construction and use of indices as potentially fascinating sources of research materials. I ought to stress, however, that other varieties of sociologists will by no means have the same reaction.'

Other major points which arose concerning the validity of our ideas concerned the concepts of 'perspective' and 'knowledge generation'. Our most extended discussion of these notions was with a philosopher whom we consulted. The following extracts are drawn from the correspondence which developed.

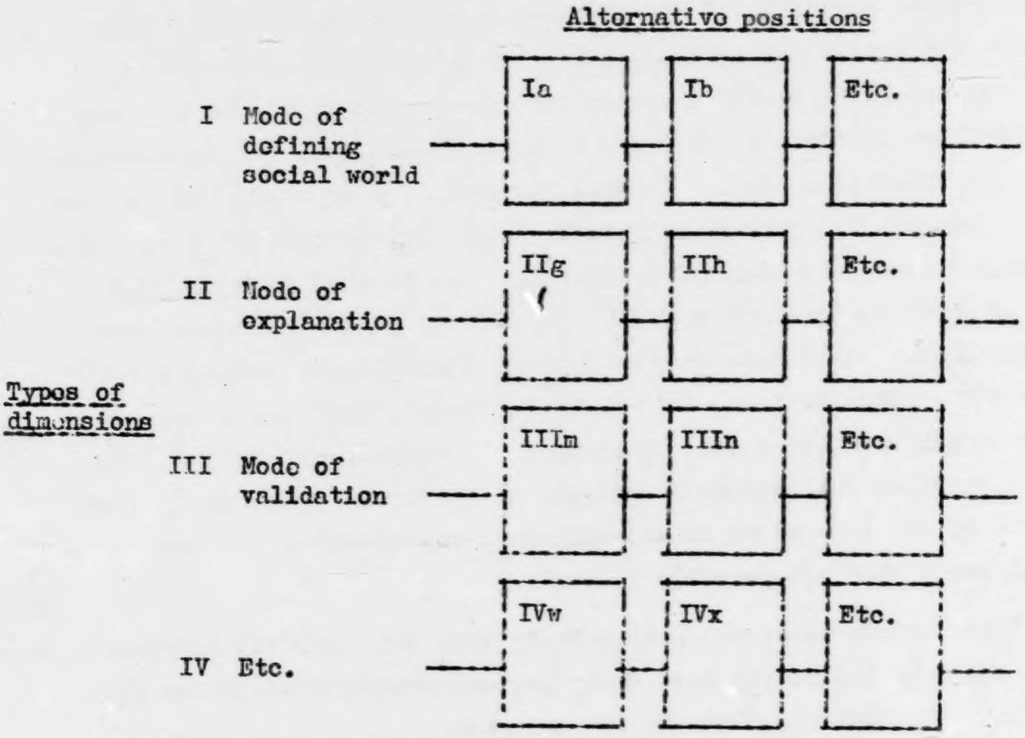
Whilst accepting the necessary variability of conceptual schemes, our correspondent noted: 'Your notion of "perspective" I find elusive. Clearly an author has a conceptual scheme, methods he employs, etc., which will place him in a tradition of some sort. But though one can get at the categories and concepts of authors in this way, I do not see how this produces a classification scheme for you.'

We were by this time planning an empirical investigation of perspectives, and the Director of the Project replied describing our ideas as follows:

'It is a fact of sociological life, with which our system must come to terms, that the variability in conceptual schemes derives in large part from fundamental differences about issues such as the nature of the subject matter with which sociology deals. Depending upon one's awareness of an individual's position on such issues, so one's interpretation of his work will often differ dramatically, and hence one's perception of its relations with other bodies of work.

'This is basically what we are talking about when we refer to people's "perspectives". However, I can well understand that you find the notion elusive, and we are at present engaged in an exercise to clarify exactly what we shall use the term to mean. If the notion is either false or meaningless, this exercise will show it up, I hope. Since we do not want to beg the question by legislating a definition in advance, we are approaching the task of definition by analysis of different kinds of sociological work which, loosely or variously, may be characterised as employing a specific perspective. For the time being, let me try to explain our notion of perspective by indicating what we expect its components to be. Briefly, we expect to identify, in the documents we analyse, a range of dimensions, both philosophical (however misused!) and sociological, on which authors tend to "locate" themselves. This will mean that we can specify the perspective of any given document in terms of a combination of positions adopted on a range of such dimensions. I do not know whether a diagram helps or not, but for what it is worth....

Framework for analysis of documents/research problems



- -

Rules for valid combination

- -

Framework for system structure: alternative perspectives

Perspective A might combine values Ia/IIh/IIIn
Perspective B might combine values Ib/IIIn Etc.

(Each set of material sharing a given perspective should be subdivided in terms of the conceptual and/or empirical problems on which the perspective tends to focus.)

'At the present stage our aim is simply to characterise the perspectives employed in the literature. We do recognise that users may employ additional perspectives, but until we have developed a way of "talking about" perspectives it is hard to see what specific questions we could put to users. However, clearly, at a later stage, we shall have to investigate, as an empirical question, the "configurations" of users' perspectives, and it could well be that this will influence the organisation of material in the eventual system, although I would hope that simply by analysis of documents we could usefully characterise a range of perspectives which have some consonance with the generally held "isms".'

Our correspondent agreed that the nature of perspectives, as we defined them, was a matter for investigation, but anticipated that the number of perspectives which were possible in logic would be too small to afford a workable framework for a classification scheme:

'All power to you in your attempts to clarify them by the procedure you outline. I do, however, have considerable misgivings though your system may in fact work. My doubts arise from the fact that you are working with a system that assumes at least the relative independence of modes of defining a social world, modes of explanation, modes of validation, etc. In fact, however, I have a horrible feeling that a mode of definition is logically very tightly related to a mode of explanation and a mode of validation. It is thus just not true that Ia could be combined with IIh and IIIn or possibly instead with IIg and IIIm. Certainly certain logical combinations here will be ruled out. What bothers me is that the number ruled out in any one instance may make the system unworkable. However, you can but try. Only a detailed philosophical analysis of particular cases could show whether or not I am right in thinking that the number of logical combinations that are possible is small.'

Since we are concerned particularly with the empirical perspectives, and only indirectly with their logical characteristics, we did not feel that this argument was too damaging.

With regard to the concept of 'knowledge generation', our philosopher correspondent questioned that a retrieval system 'can in any way contribute centrally to the generation of new knowledge'. His point was this:

'Certainly the retrieval of information results in a person acquiring new knowledge, but that is very different from the generation of knowledge, i.e., the development of new knowledge. ... New knowledge is generated by the use of what is retrieved, but by its reclassification in ways that the retrieval system cannot provide and by enquiries which it cannot suggest.'

In his reply, our Director was concerned further to clarify what we were trying to convey by the term 'knowledge generation'. He suggested that, in general, our correspondent's view of the retrieval situation represented an idealisation, whereas we were trying to come to terms with the actual situation confronting us as system designers. He identified four points at which we departed from our correspondent's analysis of what was involved.

(a) Our correspondent appeared to assume that an information system would concern itself with knowledge in the sense of facts or true propositions. Our Director stressed that, by contrast, 'we take knowledge for our practical purposes to be what passes for knowledge in the context of user activity, and this may or may not satisfy logical criteria for the use of the term, for instance, that knowledge is only contained in statements that are true'. Our respondent accepted the distinction for a retrieval system, but questioned its appropriateness as a conception of knowledge generation: '...when it comes to knowledge generation, I would have thought you were concerned with the generation of knowledge in the narrower sense'.

(b) Our Director questioned our correspondent's distinction between retrieval of information and acquisition of 'new' knowledge by a process of reclassification of what is retrieved:

'I would like to suggest that the sifting and sorting of documents is an act of classification or of imposing structure, and hence in itself a source of ideas'. Thus, 'whilst use of the literature and the generation of knowledge (ie classification or imposing structure) are analytically distinct, for all practical purposes we must regard them as often concurrent processes'. The distinction between knowledge and 'new knowledge' also seemed irrelevant to the discussion: 'knowledge is what we see when we freeze the knowledge generation process at any point in time and look at the structure'. All knowledge is thus in a sense new knowledge since 'structure changes between any one point in time and any other'.

In reply, our correspondent agreed that the sifting and sorting of documents is an act of classification or imposing structure which is a source of ideas, and went on: 'What however worries me at this point is who is doing the imposing of the structure? If the retrieval system itself is organised to impose a structure then it seems to me it can very much get in the way of valuable research.' On the question of new knowledge, our correspondent interpreted our rejection of the notion of 'newness' as irrelevant in an unintended way, to mean that we were concerning ourselves with organisation for the purpose of teaching rather than research. We felt that our arguments concerning the process of knowledge generation applied equally to course construction or to research.

(c) This led our correspondent on to comment on a further point of ours. Our Director had argued: 'Classification or imposition of structure upon a body of documents in terms of their contents is, in the broad sense in which I am using the phrase, inextricable from the generation of knowledge. ... To the extent that people consult the literature via the system in the search for knowledge, the system qua system will have created certain avenues of approach for them and in so doing will have closed others, equally valid, which users might have considered more appropriate given the choice. This is to preempt some element of the user's intellectual activity, not to perform some different and preliminary task.' He went on to suggest that there is a point to which users will accept this in the interests of saving time. Beyond this, the system 'is not simply, in any but the most trivial sense, a corpus of documents organised for retrieval as a means to an end' but rather 'the users individually working on their material, and collectively developing as a discipline, to which we, as proxies, paradoxically contribute most in so far as we do not contribute'.

Our correspondent agreed, but went further to suggest that, where classification for purposes of research was concerned, attempts to help users even thus far are 'likely to hinder first class research'. However, he referred to 'a system which will be knowledge generating', thus overlooking our stress on the fact that it is users in the process of using such a system, not the system itself, who generate knowledge. This point is one of fundamental importance for us.

(d) Our Director finally made the point that, both in terms of the nature of the conceptual schemes employed by our users, and in our view of the process by which knowledge is generated, we are working not at the individual level but in sociological terms. Our correspondent did not feel able to comment on the sociological aspect of our work. Debate of all the issues is being continued in further correspondence.

(2) Responses by information scientists

A number of theoretical issues were dealt with at length in correspondence with experts in this aspect of information science. Discussion centred around the fundamental assumptions, associated with different positions, which we had characterised in our paper. The Working Party has commented on the differences in assumptions which emerged, and therefore we do not recapitulate the arguments here.

In general, information scientists' responses did not advance discussion beyond the stage to which we had taken it in our paper, and this was disappointing. Several correspondents questioned the relevance of the issues we discussed, or at least suggested we over-emphasised their importance. One example is probably sufficient to illustrate a viewpoint of this kind:

"We may legitimately adopt any theory which helps us to attain our objectives; but, in so doing, we are not obliged:-

- to accept parts of the theory which are unhelpful or irrelevant
- to commit ourselves to any particular philosophical standpoint.'

There is a fundamental conflict of values between this kind of position and our own, and we explained our reasons for rejecting this kind of position in our paper (p.132).

An allied kind of comment, which cropped up several times, suggested that even though we may design information systems to meet users' needs, the mechanics of such systems are inherently such that special skills are required to use them. This was taken to mean that users must be helped to use the systems. Information officers were mentioned in this context. More curtly, one respondent noted that some information scientists felt that:

'...it was not wholly unreasonable to expect social scientists to accept, as have other scientists, the need for basic training in the use and handling of information retrieval and supply systems'. This was surprising since the whole burden of our paper was to argue that it is not users but existing systems that are inappropriate, and to explicate an alternative conception of retrieval as an interactive process.

Finally, we may mention comments as to the viability of what we propose. These comments tended to refer to the practical rather than the intellectual viability of realising our ideas. For example, one point was that:

'Any attempt, in a working retrieval system, accurately to represent those nuances of expression and refinements of thought which are undoubtedly important at the observer's level would tend to impair rather than improve the system.'
Several points referred to problems of responding to changes in thinking in the field, and its implications for an information system. Economic constraints were also stressed.

In so far as is possible with an idea at the blueprint stage, we had discussed questions of practically viability in our paper (p.173 et seq), in the light of our experience of running abstracting services. There was conflict of values even on those issues, because prior assumptions regarding criteria of system performance determine whether a given factor is defined as a constraint (eg any attempt at representation of refinements of thought will tend to impair retrieval) or as a variable to be manipulated according to circumstances (where users attach importance to refinement of thought, a greater degree of sensitivity to refinements of thought will be required for effective retrieval than when they do not). In the same way, observers will attach different definitions to economic factors and arrive at different assessments of viability in this respect. Comments received were less than helpful in the sense that they appeared to assess the validity of our conclusions without regard to the assumptions on which they are based. On the other hand, they were valuable as reassurance that we had not overlooked relevant variables.

III 3. (iv) [Outline of current viewpoints, Easter 1973:] The classification, for retrieval, of social science information: conference sponsored by the Office for Scientific and Technical Information

[Extract from papers circulated to those invited to attend, outlining the ground it was hoped to cover.]

Viewpoints on the social sciences

1. On the basis of the findings of his INFROSS research (Investigation into Information Requirements of the Social Sciences), Line has suggested that the social sciences fall at the soft end of a 'soft/hard' continuum, with sociology perhaps one of the softest of all! One effect of softness is a diversity of schemes of intellectual organisation and of terminology in retrieval tools, not merely across, but also within the same subject field. This is confusing, and the tools are greatly under-used. In the absence of a sound theoretical and experimental basis for planning, there is a good case for more standardisation. Even so, with a complex body of knowledge to be handled, and users with limited time and know-how, the tools may still not be used. A good deal of the findings of INFROSS seem to point strongly to the need for personal intermediaries.

2. It is alternatively possible to suggest, as the OSTI/Sociology of Education working party have done in their pre-conference paper, that current approaches to the design of information systems depend on the assumption that knowledge possesses criteria of meaningfulness which are independent of the knower. This is to transfer from the physical sciences an epistemological view which is by no means accepted by all social scientists. Recent thinking in the sociology of knowledge suggests rather that different theoretical perspectives are maintained implicitly by different academic communities whose members have learned a particular way of looking at the world, and whose knowledge takes its meaning from the social context in which it was generated. In other words, knowledge is socially constructed and socially distributed. On this view, developments in information processing which have the effect of divorcing knowledge from its context (whether defined in terms of documents or people) would seem to be dysfunctional.

These summaries draw upon the following papers:

Line, Maurice B. The information uses and needs of social scientists: an overview of INFROSS, in _Aslib Proceedings_, 1971, 23(8), 412-434.

Line, Maurice B. On the design of information systems for human beings, in _Aslib Proceedings_, 1970, 22(7), 320-335.

Contrasting rationales for the processing of social science information, I.

1. Foskett has often stressed the need for librarians and information scientists to be aware of how their clientele look at their subjects. He regards such awareness as crucial in assessing the 'pertinence' of information, that is its appropriateness to the needs of the individual user. But in the compilation of information tools, the compiler is concerned with the 'relevance'

of information. This means that he must look not to differences amongst individuals but to the consensus for guidance. Academics make their knowledge publicly available so as to make possible the process of cross-checking of results and correction of errors. Such public knowledge must be the description of a reality existing 'out there', if there is any point in communicating one's ideas of it. The structure of ideas which resides in the corpus of literature constitutes a representation of that reality or, more strictly, our understanding of it in terms of the currently accepted paradigm of reality. The intellectual organisation of material for material should be guided by the same paradigm.

2. Austin started from a position similar to that of Foskett in undertaking research for the Classification Research Group (of which both are members) into the development of a general classification scheme. In hindsight, he has argued that it is a mistake to try too hard to instil respectability into such researches by setting them into a philosophical framework. General rules for information processing are required, in the sense that the rules should hold good for subjects right across the spectrum. Different theories have useful things to say about how material may best be organised for retrieval. But one need not be bound to a single philosophical standpoint, and one need take from different theories just as much as suits the practical purpose in hand. The information scientist is concerned with symbols in documents which he extracts and uses as retrieval keys. These symbols have a relation with images in individuals' minds, and with the phenomena to which these images relate, but the images and the phenomena are not his concern. This is because, in Austin's view, they have little or no bearing on the performance of a retrieval system.

These summaries draw upon the following papers:

Foskett, D.J. Informatics, in Journal of Documentation, 1970, 26(4), 340-369.

Austin, Derek. Comments appended to Huckaby, Sarah Ann Scott. An enquiry into the theory of integrative levels as the basis for a generalized classification scheme, in Journal of Documentation, 1972, 28(2), 105-6.

Contrasting rationales for the processing of social science information, II.

3. Oldman, as a contributor to the OSTI/Sociology of Education working party pre-conference paper, analyses the academic's reasons for searching the literature as: (1) to locate data which he will reinterpret within the framework of his own ideas; (2) to locate writers whose framework of ideas, irrespective of the physical settings they investigate, is compatible with his own. Such frameworks differ across communities of academics within a discipline. These differences are crucial in terms of bringing relevant (in Foskett's sense) documents together in a retrieval system. The intellectual structure imposed by a retrieval system should thus be defined in terms of the range of dimensions comprised by different frameworks or, in other words, the dimensions by which knowledge comes to be structured. However, the way of viewing the social world which characterises a particular community develops in the course of communication within the group. Hence it rests in part on shared understandings which remain implicit. Moreover, ideas change as the discipline develops and as the social world itself changes. The degree of structure to be imposed should not therefore be great. On the one hand, users' implicit understandings must be allowed to shape the search process and define relevant categories, once a body of material broadly appropriate (in terms of the dimensions it comprises) has been identified. On the other, search patterns will necessarily change over time as boundaries between intellectual communities (writers or users) shift, and such changes must be accommodated without constant disruption of overall structure.

4. Swift, both as subject specialist and head of a unit producing three abstracting services, is in a position to consider the feasibility of implementing, in an operational service, a scheme for intellectual organisation which would take account of the considerations outlined by Oldman. Some issues for consideration are how such a scheme could be developed, how a service implementing such a scheme could be organised, what kind of collaboration between practising subject experts and information scientists would be most fruitful, and what kind of interchange of information with other services would be possible.

III Selected papers

4. Selection procedures

The first two papers are drawn from a report prepared for OSTI in December 1971, and deal with selection of journals and items respectively. The third paper is a progress report bringing the account of our work in this area up to date, and outlining some of the questions for investigation opened up by this work.

III 4. (i) Journal selection: exercise carried out end 1970/ early 1971

In selecting a sample of material for study from the viewpoint of problems for indexing and classification, it was decided to concentrate upon the journal literature. No objective way of sampling 'problems' could be devised, and we relied on the help of subject experts in constructing a relevant sample.

As a first stage, 320 journals were considered by the researchers. The researchers (1) listed around a hundred journals which form a 'core', (2) noted additional journals which contain further material to be studied if time permitted, and (3) identified those to which no further attention need be paid. (See appendix to this paper for details of journals.)

The sample of 100 journals included examples of all kinds of materials which sociologists of education may need to consult for their work. It was assumed that they would approach the material sociologically, even if their main interest is in the relevance of sociological studies to educational problems rather than in sociology as sociology.

Selection was made from three sources: Sociology of Education Abstracts (SEA), additional social science journals covered by Sociological Abstracts (SA) and additional educational journals held by the London Institute of Education Library. In the case of journals identified through SEA and SA, originals were examined wherever possible. Issues for 1968-70 were studied.

Items in each journals were considered from the following points of view:

- Proportion of relevant material
- Nature of relevance:
  - Sociological (including social psychological) analyses of education (broadly defined)
  - Sociological writing having implications for the sociological study of education
  - Educational writing having implications for the sociological study of education
  - Centrality of topics dealt with
  - 'Substantiality' of treatment

With regard to the proportion of relevant material, in general only journals containing at least one or two items per year of a sociological nature (whether directly bearing on education or not) were included in our first category. In the case of educational journals, 'sociological' was interpreted broadly to include 'social' but the 'substantiality' criterion was defined so as to exclude most purely descriptive material.

Certain whole areas of study were excluded as marginal (eg eugenics), also others covered by other services (crime and delinquency). The criteria were relaxed a little in other areas which it was felt were important and should be represented, but this was not taken too far - thus several journals dealing with technical education and industrial training were excluded but the British Journal of Industrial Relations was included. In borderline areas, where there were several relevant specialised journals, the cluster was excluded or included as a cluster; where there was an isolated journal only it was normally excluded. With regard to methodology, journals with a broad scope were included but not anything as restricted in coverage as, for instance, Multivariate Behavioral Research.

As far as possible, specific journals were taken as reference points in order to compare across both subject areas and across journals rather than comparing items only with items in the same or similar journals. For instance 'popular' journals, if relevant, were included or excluded on the basis of comparison with a journal such as New Society which was felt to contain a range of writing which sociologists of education might well wish to refer.

The second category (to be studied if time permitted) contained two kinds of journal. First, there were those containing items which qualify on all criteria but there were too few items for our purposes. Additionally, there were journals containing a reasonable amount of material which was relevant in subject matter but too high a proportion of this, for our purposes, was non-sociological or non-substantial by comparison with journals in the first hundred.

The third category contained those journals to which we felt the sociologist of education would not find it useful to be referred.

As a second stage, we asked our Working Party of subject experts to draw on their knowledge of the journal literature, and to look at our first category of material from the viewpoint of similarities and differences in the nature of the problems addressed. Since we wished to index material in such a way as to bring out the distinctive character of a given item, we anticipated that differences in nature of problem addressed would be a major factor in problems for indexing. Our Working Party identified about 40 journals likely to raise problems for indexing defined in this way. The research team then divided this set of journals into two subsets: those with wide coverage and those focussing on a particular area of study (eg Administrative Science Quarterly).

Our plan was successively to study the indexing problems represented by these two subsets of journal material. We intended then to apply the experience we had gained to the processing of reports and books, in terms of either 'whole item' treatment (where a single investigation is reported or a single theme is developed), or 'chapter by chapter' treatment (to be used wherever possible, ie when individual chapters are relatively self contained).*

---

* In the event, pressure of other work has prevented us from making a detailed study of the latter subset of journals, or of books, up to this time (June, 1973).

Appendix to
Paper III 4. (i)

LIST OF JOURNALS FROM WHICH SAMPLE WAS DRAWN

Category I

✓ Administrative Science Quarterly
Administrators' Notebook
Adult Education (U.S.A.)
American Behavioural Scientist
✓✓ American Journal of Sociology
✓ American Political Science Review
✓✓ American Sociological Review
Annals of the American Academy of Political & Social Science
Behavioural Science
✓✓ British Journal of Educational Psychology
✓✓ British Journal of Educational Studies
✓ British Journal of Industrial Relations
British Journal of Social & Clinical Psychology
✓✓ British Journal of Sociology
California Journal of Educational Research
✓✓ Child Development
Child Development Monographs
✓✓ Comparative Education
Comparative Education Review
✓✓ Daedalus
Durham Research Review
✓ Educational Administration Quarterly
Educational Philosophy and Theory
✓✓ Educational Research
Educational Theory
✓✓ Education and Urban Society
Genetic Psychology Monographs
✓✓ Harvard Educational Review
Human Development
✓ Human Organisation
Human Relations
Industrial and Labour Relations Review
International Journal of Educational Sciences
Journal of American Statistical Association
✓ Journal of Applied Behavioural Science
✓ Journal of Biosocial Science
Journal of Counselling Psychology
Journal of Creative Behaviour
✓ Journal of Curriculum Studies
Journal of Educational Psychology
Journal of Educational Research
✓ Journal of Educational Thought
✓ Journal of Experimental Education
Journal of Genetic Psychology
✓ Journal of Health & Social Behaviour
Journal of Management Studies
✓ Journal of Marriage & the Family
Journal of Occupational Psychology
Journal of Personality
Journal of Political Economy
✓ Journal of Royal Statistical Society (Series A)
✓✓ Journal of Social Issues
✓ Journal of Social Psychology
Language and Speech
✓ Merrill Palmer Quarterly
Minerva
Moral Education
✓✓ New Society
Pacific Sociological Review
✓ Personnel & Guidance Journal
✓ Personnel Management
Political Science Quarterly
Proceedings of the Academy of Political Science
Record
Research in Education
Review of Educational Research
Rural Sociology
✓✓ School Review
Science
Social & Economic Administration
✓ Social Forces
✓ Social Problems
Social Science & Medicine
✓ Socio-economic Planning Sciences
Sociological Analysis
Sociological Bulletin
Sociological Inquiry
✓✓ Sociological Review
✓✓ Sociology
✓ Sociology and Social Research
✓✓ Sociology of Education
✓ Sociometry
Soviet Education
✓ Soviet Sociology
Technology and Society
✓✓ Universities Quarterly
Urban Review
Vocational Guidance Quarterly

N.B. Items ticked indicate sample selected by Working Party (double ticking refers to subset studied by researchers).

LIST OF JOURNALS (continued)

Category 2

A.A.U.P. Bulletin
AV Communications Review
Adolescence
American Economic Review
American Journal of Orthopsychiatry
Developmental Psychology
British Journal of Medical Education
Change in Higher Education
✓Education for Teaching
Educational Forum
Educational Record
Educational Review
Family Coordinator
The Human Context
Interchange
International Development Review
Irish Journal of Education
Journal of Aesthetic Education
✓Journal of Educational Technology
Journal of Communication
Journal of Conflict Resolution
Journal of Educational Measurement
Journal of Human Relations
✓Journal of Human Resources
Journal of Law and Economics
Journal of Negro Education
✓Journal of R & D in Education
Journal of Social Work Process
Journal of Teacher Education in U.S.A.
Junior College Journal
Kansas Studies in Education
Law and Society Review
Management International Review
Management Science
Measurement and Evaluation in Guidance
Midwest Journal of Political Science
National Elementary Principal
NEA Research Bulletin
Past and Present
Personnel Administration
Personnel and Training Management
Personnel Psychology
Phylon
Psychology in the Schools
Religious Education
Rocky Mountain Social Science Journal
School and Society
School Counsellor
Scottish Educational Studies
Social Compass
Studies in Adult Education
Teacher Education
Transaction
✓Trends in Education
Urban Education
Vocational Aspect
Vocational Guidance Quarterly

LIST OF JOURNALS (continued)

Category 3

Adult Education (U.K.)
Adult Leadership
American Anthropologist
American Council of Learned Societies Newsletter
American Education Research Journal
American Journal of Economics & Sociology
American Journal of Mental Deficiency
American Scholar
American School and University
American Scientist
Antioch Review
Arts in Society
Assignement Children
B.A.C.I.E. Journal
B.A.C.I.E. News
Buffalo Law Review
Bulletin of N.A.S.S.P.
Bulletin of University of London Institute of Education
California Institute of Education Bulletin
Cambridge Institute of Education Bulletin
Catholic Education Review
Centennial Review
Child Care
Child Welfare
Christian Century
Clearing House for Social Literature
Commentary
Community Development Journal
Community Mental Health Journal
Comparative Studies in Society and History
Cornell Journal of Social Relations
Dimensions
Dissent
Education and Social Science
Educational Panorama
✓Educational Leadership
Elementary School Journal
Elementary School Principal
Ethnology
Eugenics Quarterly
Eugenics Review
European Teacher
Exceptional Children
Explorations in Entrepreneurial History
Economic Development and Cultural Change
Education and Training
Florida Educational Research and Development Council Research Bulletin
Fordham Law Review
Froebel Journal
Forum
Forward Trends
Further Education
Futures
George Washington Law Review
The Gifted Child Quarterly
Harvard Graduate School of Education Bulletin
Howard Journal of Penology
Head Teachers Review
Higher Education Journal
Higher Education Review
High School Journal
Human Biology
Industrial and Commercial Training
Industrial Relations
Industrial Society
Industrial Training International
Improving College and University Teaching
International Journal of Electrical Engineering Education
International Journal of Offender Therapy
International Journal of Social Psychiatry
International Migration Review
Journal of Advertising Research
Journal of American History
Journal of American Statistical Association
Journal of Applied Behavioural Analysis
Journal of the Association of Teachers of Management
Journal of Broadcasting
Journal of Child Psychology and Psychiatry
Journal of Criminal Law/Criminology and Police Science
Journal of C.R.A.C.
Journal of Education
Journal of Educational Administration and History
Journal of Engineering Education
Journal of Family Law
Journal of General Education
Journal of Higher Education
Journal of History of the Behavioural Sciences
Journal of Individual Psychology
Journal of Industrial and Technical Education
Journal of Industrial Relations
Journal of Research in Crime and Delinquency
Journal of Sex Research
Journal of Social History
Journal of Special Education
Journalism Quarterly
Jurimetrics Journal
Law and Contemporary Problems
Man
Management of Personnel Quarterly
Massachusettes Review
Maryland Law Review
Mental Health
Mental Hygiene

LIST OF JOURNALS (continued)

Category 3 (continued)

Mental Retardation
Michigan State University Quarterly Bulletin
Midwest Journal of Political Science
Midwest Quarterly
Millbank Memorial Fund Quarterly
Mississippi Quarterly
Multivariate Behavioural Research
New Era
New University (formerly New Education)
North Western Law Review
Partisan Review
Peabody Journal of Education
Population Studies
Psychological Bulletin
Psychological Reports
Psychological Review
Phalanx
Phi-Delta Kappan
Philosophy & Phenomenological Research
Philosophy of Science
Psychiatry
Public Interest
Public Opinion Quarterly
Quarterly Journal of Studies on Alcohol
Race
Race Today
Research Reports in Social Science
Review of Religious Research
Russian Review
Science and Society
Science Forum
Scottish Education Journal
Screen
Smith College Studies in Social Work
Social Biology
Social Casework
Social Education
Social Science
Social Science Quarterly
Social Service Review
Social Work
Southwestern Journal of Anthropology
Southwestern Law Journal
Soviet Anthropology and Archaeology
Soviet Review
Special Education
Studies in Comparative International Development
Studies in Philosophy of Education
Summation
Teaching
Technical Journal
Technology and Culture
Theory into Practice
Todays Education
Transaction
University of Chicago Law Review
University of Toronto Law Journal
University of Detroit Journal of Urban Law
University of Maine Law Review
University of Manchester School of Education Gazette
University of Washington Journal of Sociology
Urban Affairs Quarterly
Urban Studies
Villanova Law Review
Welfare in Review
Well Being
Whore
Wisconsin Law Review
Wisconsin Sociologist
Yale Review
Young Children

III 4. (ii) Item selection: development of procedures January - June 1971

The researchers, concentrating first on sociological journals amongst those starred on the lists appended to the previous paper, were able to identify a few items which clearly had no relevance for any sociologist of education, together with a number of items which clearly did have relevance for all sociologists of education. The problem lay in a broad band of material which was in some sense relevant yet would, if all included, constitute something closer to a bibliography in sociology than a bibliography for the sociology of education.

It was pointed out that it would be true to say that all sociology was relevant to the study of education. Beyond the identification of sociological writing and research that dealt specifically with educational subjects or settings, selection from other sociological work would run the risk of individual and perhaps idiosyncratic judgments of relevance.

This view can be challenged on two grounds. First, not all sociological work in which educational subjects or settings feature is necessarily relevant to the study of the sociology of education. A researcher might happen to have used a sample of students merely as convenient 'guinea-pigs' and not by virtue of interest in them as students. In this case, the problem studied will not necessarily be one of concern to the sociologist of education.

Secondly, some sociological work can be said to be 'more' relevant than some other to the study of the sociology of education. Some studies are such that a similar study could be carried out in relation to some educational problem or problems without modification of approach, whilst in others considerable modification would be required. This gave rise to the notion of 'first order' relevance as compared with second order of remove of relevance. Such judgments involve the exercise of subject knowledge, but the principle is one which can be applied with a reasonable degree of reliability.

The 'order of remove' principle provided an acceptable working basis for selection of sociological material, both empirical studies and middle range and context specific theory. A somewhat different principle was required to handle work treating 'pure' theory. In this case, a distinction could be made between comment on or working with established theory and resynthesis or new theory, the latter being of immediate interest to the sociologist of education.

These principles emerged after detailed discussion of 'borderline' items, in an attempt to identify the determining factors in the general 'feeling' that an item should or should not be included as relevant to 'doing' the sociology of education, and to study the appropriateness of alternative grounds for decisions. An interesting point in this exercise was a growing mutual understanding amongst educationalist and sociologist members of the Working Party of each others' viewpoints, which enabled them to discuss particular items effectively yet without compromise of viewpoints.

A similar process was adopted in the case of non-sociological material and further broadly sociological material. The main problem with non-sociological material was the vast amount of material which in some sense may represent raw data for sociological study. Some of this work is theory-based and may even (eg some economics or social psychology) be claimed by some sociologists as sociology. The range extends to the articles of faith and rules of thumb which form a large part of the contents of, for instance, 'trade journals' for educationalists, and pure journalism. Much of this material is clearly raw data for the future social historian rather than for the contemporary sociologist of education.

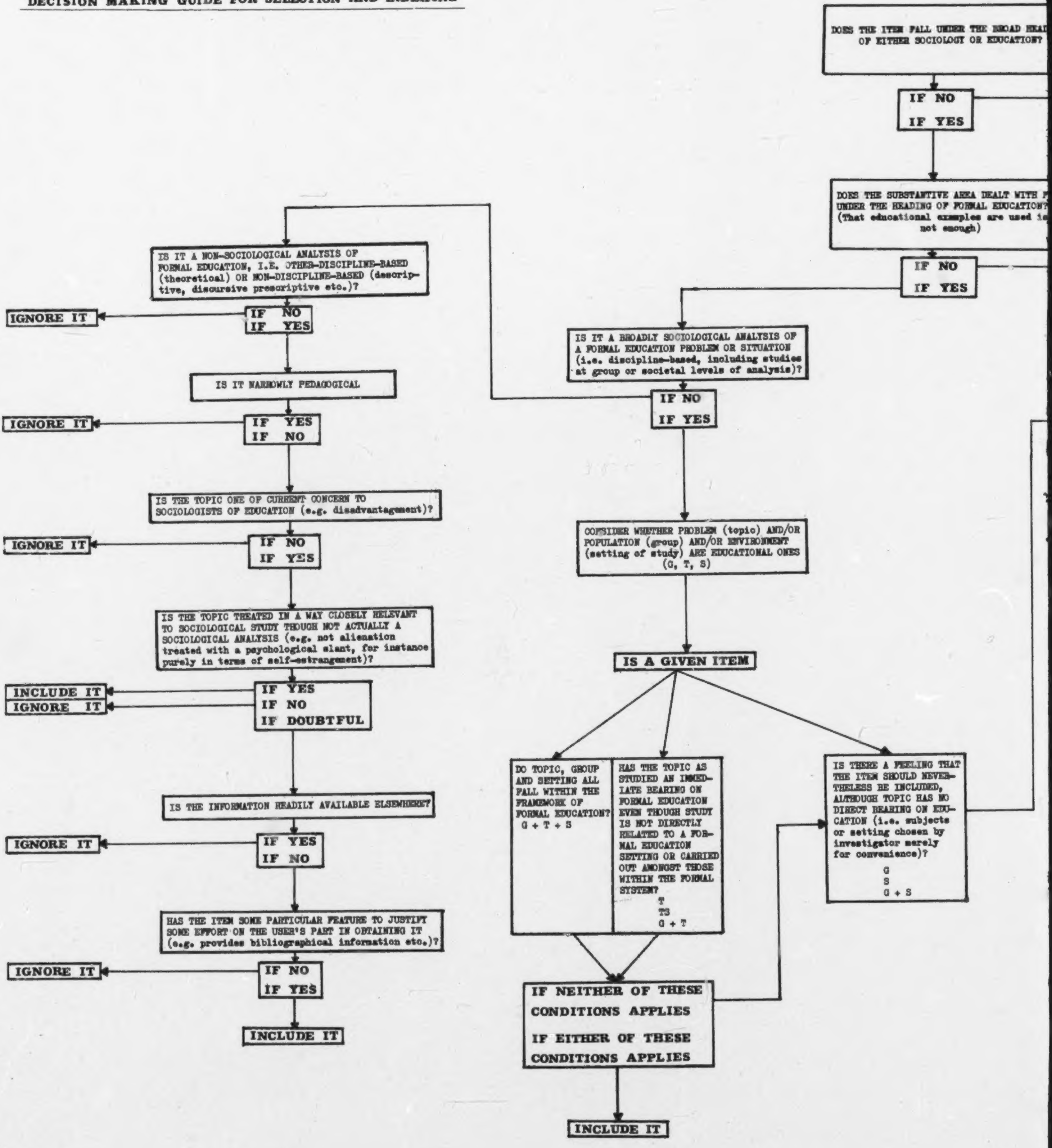
It became clear that whilst the aim in the case of sociological material was to identify grounds for inclusion, in the case of non-sociological material the range was so great and so varied that selection principles should be on the whole exclusive rather than inclusive. A number of grounds for exclusion were identified such that, even if an item failed on only one count, it would be excluded.

This work was developed into a provisional flow chart of the decision making process. (See Appendix A to this paper.) It should be stressed that such an aid to selection does not of itself ensure consistency of selection, even while it may considerably reduce inconsistency. Preliminary comparison of items provisionally included with those provisionally excluded indicated that cut off points may shift even when the criteria remain constant, but reviews of this kind can be used to establish precedents to guide future decisions on items falling just at a borderline, and possibly to refine the definitions of criteria.

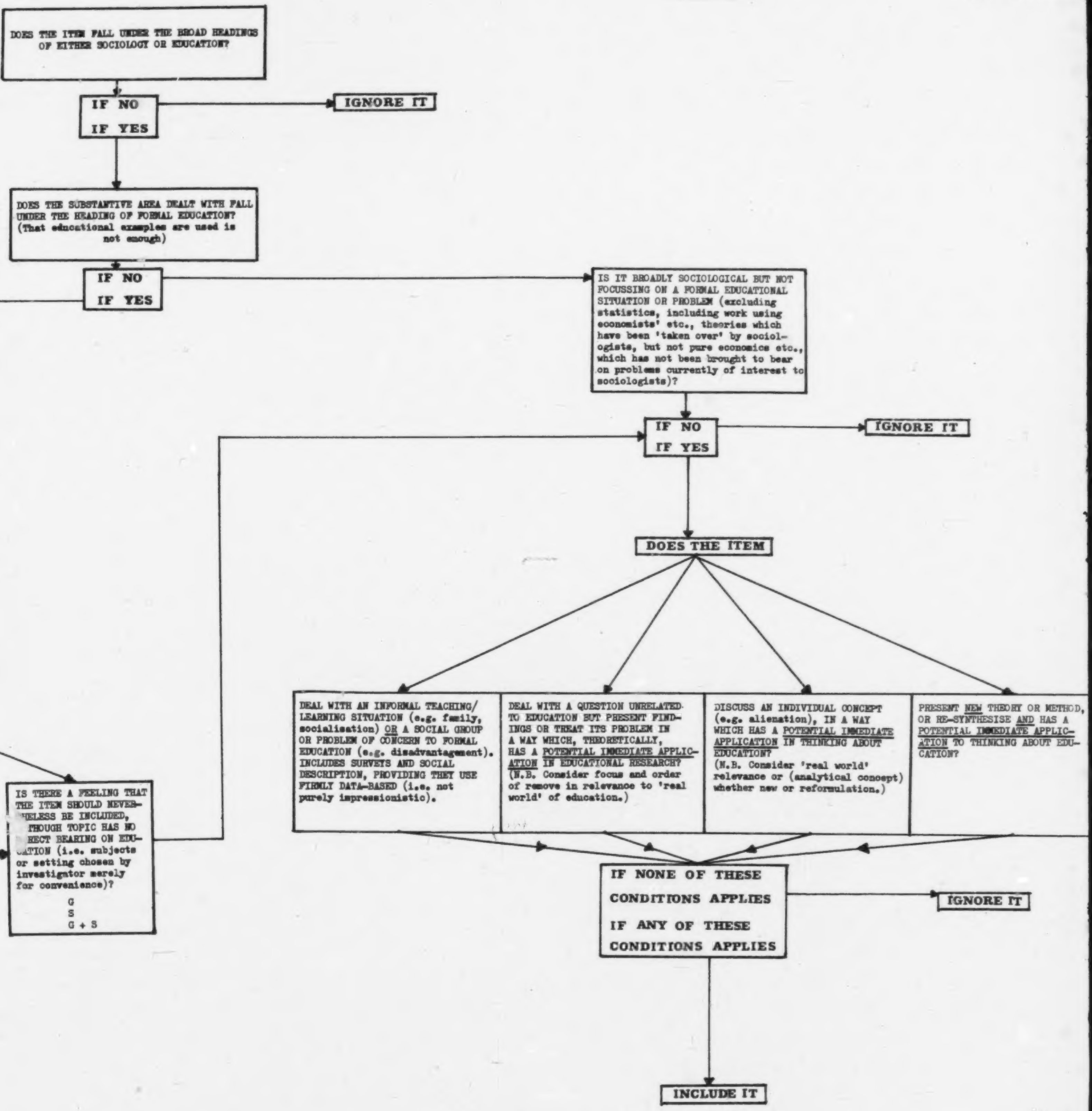
The effect is to develop a case law to guide implementation of the formal guidelines for decision making. Appendix B of this paper exemplifies an emerging case law. The advantage of this approach is that the case law may be adapted in relation to the current output of documents (eg criteria for a given area of investigation may become progressively more rigorously applied from the time when it represents a growth point to the time when a 'flood gates' situation obtains), whilst the general principles remain constant.

OSTI/SEA PROJECT

DECISION MAKING GUIDE FOR SELECTION AND INDEXING


DOES THE ITEM FALL UNDER THE BROAD HEA
OF EITHER SOCIOLOGY OR EDUCATION?
IF NO
IF YES
DOES THE SUBSTANTIVE AREA DEALT WITH F
UNDER THE HEADING OF FORMAL EDUCATION?
(That educational examples are used i
not enough)
IF NO
IF YES
IS IT A BROADLY SOCIOLOGICAL ANALYSIS OF A FORMAL EDUCATION PROBLEM OR SITUATION (i.e. discipline-based, including studies at group or societal levels of analysis)?
IF NO
IF YES
IS IT A NON-SOCIOLOGICAL ANALYSIS OF FORMAL EDUCATION, I.E. OTHER-DISCIPLINE-BASED (theoretical) OR NON-DISCIPLINE-BASED (descriptive, discursive prescriptive etc.)?
IGNORE IT
IF NO
IF YES
IS IT NARROWLY PEDAGOGICAL
IGNORE IT
IF YES
IF NO
IS THE TOPIC ONE OF CURRENT CONCERN TO SOCIOLOGISTS OF EDUCATION (e.g. disadvantagement)?
IGNORE IT
IF NO
IF YES
IS THE TOPIC TREATED IN A WAY CLOSELY RELEVANT TO SOCIOLOGICAL STUDY THOUGH NOT ACTUALLY A SOCIOLOGICAL ANALYSIS (e.g. not alienation treated with a psychological slant, for instance purely in terms of self-estrangement)?
INCLUDE IT
IGNORE IT
IF YES
IF NO
IF DOUBTFUL
IS THE INFORMATION READILY AVAILABLE ELSEWHERE?
IGNORE IT
IF YES
IF NO
HAS THE ITEM SOME PARTICULAR FEATURE TO JUSTIFY SOME EFFORT ON THE USER'S PART IN OBTAINING IT (e.g. provides bibliographical information etc.)?
IGNORE IT
IF NO
IF YES
INCLUDE IT
CONSIDER WHETHER PROBLEM (topic) AND/OR POPULATION (group) AND/OR ENVIRONMENT (setting of study) ARE EDUCATIONAL ONES (G, T, S)
IS A GIVEN ITEM
DO TOPIC, GROUP AND SETTING ALL FALL WITHIN THE FRAMEWORK OF FORMAL EDUCATION?
G + T + S
HAS THE TOPIC AS STUDIED AN IMMEDIATE BEARING ON FORMAL EDUCATION EVEN THOUGH STUDY IS NOT DIRECTLY RELATED TO A FORMAL EDUCATION SETTING OR CARRIED OUT AMONGST THOSE WITHIN THE FORMAL SYSTEM?
T
TS
G + T
IS THERE A FEELING THAT THE ITEM SHOULD NEVERTHELESS BE INCLUDED, ALTHOUGH TOPIC HAS NO DIRECT BEARING ON EDUCATION (i.e. subjects or setting chosen by investigator merely for convenience)?
G
S
G + S
IF NEITHER OF THESE CONDITIONS APPLIES
IF EITHER OF THESE CONDITIONS APPLIES
INCLUDE IT

APPENDIX A


DOES THE ITEM FALL UNDER THE BROAD HEADINGS OF EITHER SOCIOLOGY OR EDUCATION?
IF NO
IF YES
IGNORE IT
DOES THE SUBSTANTIVE AREA DEALT WITH FALL UNDER THE HEADING OF FORMAL EDUCATION? (That educational examples are used is not enough)
IF NO
IF YES
IS IT BROADLY SOCIOLOGICAL BUT NOT FOCUSSING ON A FORMAL EDUCATIONAL SITUATION OR PROBLEM (excluding statistics, including work using economists' etc., theories which have been 'taken over' by sociologists, but not pure economics etc., which has not been brought to bear on problems currently of interest to sociologists)?
IF NO
IF YES
IGNORE IT
DOES THE ITEM
DEAL WITH AN INFORMAL TEACHING/LEARNING SITUATION (e.g. family, socialisation) OR A SOCIAL GROUP OR PROBLEM OF CONCERN TO FORMAL EDUCATION (e.g. disadvantagement). INCLUDES SURVEYS AND SOCIAL DESCRIPTION, PROVIDING THEY USE FIRMLY DATA-BASED (i.e. not purely impressionistic).
DEAL WITH A QUESTION UNRELATED TO EDUCATION BUT PRESENT FINDINGS OR THEAT ITS PROBLEM IN A WAY WHICH, THEORETICALLY, HAS A POTENTIAL IMMEDIATE APPLICATION IN EDUCATIONAL RESEARCH? (N.B. Consider focus and order of remove in relevance to 'real world' of education.)
DISCUSS AN INDIVIDUAL CONCEPT (e.g. alienation), IN A WAY WHICH HAS A POTENTIAL IMMEDIATE APPLICATION IN THINKING ABOUT EDUCATION? (N.B. Consider 'real world' relevance or (analytical concept) whether new or reformulation.)
PRESENT NEW THEORY OR METHOD, OR RE-SYNTHESISE AND HAS A POTENTIAL IMMEDIATE APPLICATION TO THINKING ABOUT EDUCATION?
IS THERE A FEELING THAT THE ITEM SHOULD NEVER-HELESS BE INCLUDED, THOUGH TOPIC HAS NO RECT BEARING ON EDUCATION (i.e. subjects or setting chosen by investigator merely for convenience)?
G
S
G + S
IF NONE OF THESE CONDITIONS APPLIES
IGNORE IT
IF ANY OF THESE CONDITIONS APPLIES
INCLUDE IT

APPENDIX B to
Paper III 4. (ii)

Notes on decision making guide for selection and indexing of Material of broader sociological interest

A Areas related to education

Certain 'areas' of material are included as a matter of general principle whereas others are excluded. Within these areas further discrimination may be involved (e g. articles on the family as a social unit are considered relevant whereas items specifically on conjugal relationships are not). Further areas may be added in the future as more material is selected.

Areas related to education

- Family
- Adolescence
- Socialization
- Careers, occupations, professions
- Diffusion of knowledge

Excluded areas

- Guidance
- Deviance: crime
  - mental ill health

B Areas unrelated to education

Here we are concerned with material that uses theory in the empirical study of situations other than educational situations. To be of relevance to the study of education, such material must do more than increase understanding of the particular problem studied, and also use of concepts must be such that they are capable of application to education in a 'real world' sense.

The real world' application takes account of the social unit e.g. 'community' and 'neighbourhood' do not 'underlie' any social units within education and/or the process involved, .g social mobility as a concept does not 'underlie' educational mobility. Further if an item deals with a process or unit e.g. elite group which is relevant to education but in such a way that it is situation specific it will not be deemed relevant in this category.

N.B. Methodological aspects of such works are not taken into account under this heading.

C Discussion of specific concepts

To be accepted on these grounds an item must either: (a) deal with a concept or an aspect of a concept in a way that has a 'real world' educational application (e.g. transitory/permanence dimension of organization life span is not particularly relevant to study of organizations in education), or (b) in the case of purely analytic concepts with no direct empirical referent (e.g. status inconsistency) the concept must be a newly created one or a major reformulation of an existing one.

D Presentation of general approaches or methods

To be included under this heading an item must either present a new theory or a synthesis of existing theories (e.g. an item on functionalism, per se, would be excluded but one hypothesizing a synthesis of functionalism and conflict theories would be included).

NOTE: As a general principle we will err on the side of inclusion, and on the rare occasion where every instinct contradicts the rules, instinct will be the deciding factor.

III 4. (iii) Progress report on development of selection procedures, June 1973

Our work in the area of selection during the period since the previous report comprises a phase of development. The main task has been simply on-going selection of items from our sample, working within the framework of our selection chart, so as to develop case law. We are now approaching a point at which we have precedents for dealing with a substantial proportion of the items we are encountering in the present sample, and at which the thinking underlying these precedents is proving to have a reasonable degree of generalisability beyond the items to which decisions initially related.

We have not yet, however, reached a stage at which we can report a detailed account of our selection procedures. This report is simply intended to bring the record up to date and to describe the point our work has now reached. We would not wish to suggest that it represents any completed programme of investigation. It would be more appropriate to say that our work in this area to date has given us the means by which to investigate the problems and potential of employing an approach to selection which seeks to take explicit account of the decision making process involved. It has also given us a clearer idea of the questions on which such investigation should focus.

A review of our selection decisions to date, now in progress, is beginning to explore some of these questions. It is part of our thinking that selection procedures should above all be responsive to various and changing emphases and approaches in a field. There are, therefore, several aspects to the investigation which is in hand. We are concerned with the content of our selection decisions to date from the standpoint of their appropriateness at this point in time. This concern, however, derives not from a wish to evaluate our product from the point of view of current performance in selection, but from the aim of improving the procedures we are employing, so that they lead to decisions appropriate to this point in time. We are studying our procedures also from the point of view that, when we review subsequent decisions at some future time, these later decisions should be appropriate to that point in time. We have in mind too that, quite apart from the particular content, present or future, of ideas in our field, the procedures should be capable of being operationalised in a range of other fields of study.

The general aim of investigation is, therefore, a developmental one of improving the design of our procedures in the light of the purposes we wish them to serve. This involves working at the level both of general questions relevant in selection and characteristics of particular situations which may arise. The distinction between the selection process and particular case law(s) which may be associated with it was used in our earlier report as descriptive of our work. Perhaps the major development in our thinking since that time is that we have come to recognise the value of the notion of case law in permitting selection to be sensitive to a range of circumstances yet consistent with an overall rationale, and this has now become a guiding idea.

We briefly describe our investigation under the following headings: current formulation of case law; implementation of procedures in SEA context; wider relevance.

Current formulation of case law

The basis for this aspect of the investigation is a comparison of inclusion and exclusion decisions on items in our sample, in terms of reasons for decisions. The team has kept a record of reasons supporting each decision. Situations have arisen in the process of selection where it has been necessary to question precedents already established or to create precedents. Such situations may affect the case law. For instance, problems of slant have been found with material relating to rewards and punishments. Studies of culture contact have come to be regarded as an area 'in flood'. Additional considerations or appropriate cut off points have to be introduced in such situations. Analysis of the thinking involved enables a formulation of case law which may guide decisions when similar situations are encountered in the future, yet at the same time recognising that, in combination with different factors, a given situation may give rise to different decisions. Major changes may be required. For instance, comments from many users have indicated that an early decision to exclude general methodology on principle is inappropriate. Reversal of this decision involves an extension of case law to take account of new questions for selection raised by this material. These kinds of modifications to practice are exactly comparable to modifications which will be characteristic of the operational situation. Our purpose is to ensure that our case law is formulated in such a way as to accommodate them within the general framework provided by our chart, without discuption of it. We have been encouraged to find that the emergence in our sample of an increasing number of studies employing a phenomenological approach has required no major rethinking of our framework.

Implementation of procedures in SEA context

A comparison of the research team's selection decisions with those of SEA abstractors in respect of the team's sample of material is now in preparation. This will enable the team further to improve the thinking built in to its case law. From the SEA point of view, the team's work will provide a language in which to talk about SEA'S future selection policy. It is intended that SEA will adopt the approach, but there will be need for further case law. Extension of the team's sample of journals is envisaged as the next step to this end, and consultation with SEA abstractors is desirable so as gradually to mesh in with SEA practice. Organisational questions are also in our minds in undertaking this exercise. It will not be possible to analyse all decisions in an operational situation. Quality control measures can at best comprise continuing spot checks or occasional complete but point in time analyses. We need to determine the most effective way of monitoring the handling of areas sensitive to change and emerging areas. For this purpose we are thinking in terms of identifying strategic or borderline items which raise issues on which selection decisions turn, and which can be used, once a precedent has been established, to exemplify alternative decisions in relation to alternative circumstances. Regular consultation with abstractors and users in the field will be essential for continuing updating of case law. In the longer term, perhaps at ten year intervals, we would envisage overhaul of the thinking embodied in the general framework. Guidance on what this may involve will be obtained by analysing the selection for an early volume of SEA in relation to our selection chart.

Wider relevance

Our terms of reference - the selection of material relevant to the sociological analysis of education - means that our procedures have been designed to handle a wide range of material (discipline- and non-discipline-based) in addition to material narrowly defined as sociological studies of formal educational situations. This has been beneficial in ensuring that we have not taken an encapsulated view of our field but have constantly been reminded of points of contact with other fields. It is therefore reasonable to suppose that, even if the particular case law we are developing is primarily of domestic interest, the general framework we have devised would have wider applicability. This has yet to be put to the test. Consultation with subject experts has indicated that they find nothing unacceptable in the thinking underlying our selection chart. Preliminary discussion with non-subject experts, however, indicates a conflicting approach which seems to reflect a traditional subject-bound view of selection. Thus, instead of relating document selection to the intellectual activity it is intended to support, the subject-bound approach looks for boundaries in the subject matter under study. This leads to definitions of relevance described as high/low or target field (direct)/related field (indirect), suggesting that relevance is a unitary variable rather than a complex phenomenon, and that it is independent of situational factors. We believe this approach to derive from an over-simplified conceptualisation of the retrieval situation. To the extent that selection practices in other contexts are based on such a conceptualisation, it is unlikely that our approach will be seen to have wider applicability. However, private collections built up by subject experts, and made available to their students, will afford an opportunity to assess our ideas in the context of related specialised areas.

The development of our indexing and classification scheme may additionally influence the range of material we decide to include. Even apart from this, investigation of the questions we have discussed here will not be completed in the term of the present project. Reports will be issued as the findings of different aspects of the investigation become available.

III Selected papers

5. Representation of individual documents

Initially, our approach to the processing of individual documents was influenced by our experimentation with PRECIS and our need to have subject experts study the descriptions we produced in relation to the documents they represented. A two stage process was devised, involving preparation of a fairly full summary, which subject experts could use as a document substitute against which to assess the descriptions we produced as a second stage. We have found it helpful to retain the two stage system even though the original reason for it has gone, and the procedures associated with each stage have been modified.

The first paper outlines the thinking by which we have arrived at our rationale for processing. It involves an attention to detail that some regard as an unnecessary luxury. The second and third papers exemplify experimental work in which we tried and failed to find an acceptable short cut. The fourth paper illustrates reasons for believing that lack of attention to 'detail' impedes retrieval. The fifth and sixth papers deal individually with our two stages of processing as we have come to operationalise them. Finally, comments of authors to whom we have circulated our work for comment are reported.

III 5. (i) Development of our ideas about processing individual documents, November 1970 to date.

Our early ideas were greatly influenced by the requirement to relate our practice to working within the framework of the PRECIS indexing system, as described in an earlier paper. The PRECIS system, in using descriptions as a basis for a 'word extraction' method of generating subject headings, was found to be a major source of distortion in descriptions of individual documents. Since that time, we have been developing procedures for handling individual documents independently of considerations of the generation of subject headings, with the primary aim of producing non-distorting descriptions. We did not rule out the possibility that, given acceptable non-distorting descriptions, subject headings might be systematically but indirectly related to descriptions. Our work on this question is touched on in the papers on overall intellectual organisation.

Since our PRECIS period, our aim has been to represent documents in their own terms. The progression of our ideas concerning the description of individual documents may be summed up as a deepening understanding of what is meant by representing a document in its own terms. We have moved from a rejection of the controlled language and structure PRECIS led us to impose on our documents to natural language and a search for more appropriate structure, and thence to the notion of 'not interfering' with either language or structure of the document. As a result, our practice has changed from early attempts to decide what should be included in a description and how the elements should be related, to thinking in terms of what can be excluded from the document, a description being what remains after the least important elements and relationships have been excluded. A two stage approach to processing, originally designed for our needs when working in the PRECIS context, has become adapted to our present view of processing.

In the processing of successive sets of documents, several inescapable points emerged to lead us to this practice. We had been disturbed by the assertion of the PRECIS team that the formulation of a statement of subject, which provides the starting point for the indexer, is a professional skill which involves objective analysis of concepts and their interrelationships without reference to the content of a document. It was suggested that subject experts tend to perform badly at this task because they are too involved with the subject matter to see the 'wood for the trees'. We were well aware of the 'wood for the trees' problem but, in our experience, objective techniques did not solve it. The effect was that exercise of such techniques could lead to several alternative interpretations of a document, and an absence of any rationale to determine which, if any, represented it in its own terms. Another serious problem was that a superficially 'correct' subject statement could be highly misleading. The danger here is that to which newspaper headlines are open. For example, a headline 'Bloggstown swings to left' in an electoral campaign might conceal the fact that the elected candidate received a small proportion of the vote, but won the election by virtue of a split vote situation. We have to accept, then, that to represent a document in its own terms, techniques for preparing descriptions should be guided by understanding of the document to be described.

It was clear that this would involve a greater outlay of time per document than a more superficial scanning of a document for salient variables or summary phrases concerning theoretical framework, content and source of data, the three aspects of a document by which we came to structure our descriptions in our post-PRECIS period. OSTI questioned the need for understanding as a 'luxury' approach to processing. Experimentation with various aids and short cuts, however, failed to produce an acceptable alternative approach (two experiments are reported elsewhere in this section). We also reconsidered whether the distortion introduced by alternatives to 'understanding' would be seriously detrimental to the retrieval of relevant documents. Another paper on the mishandling of documents with which we are concerned indicates that distortion does matter.

Processing based on understanding, however, does not of itself eliminate distortion. There is a sense in which a document has as many meanings as there are people who read it. The individual cannot avoid putting a construction on what he reads. Individuals preparing descriptions will tend to employ a set pattern of response to documents, irrespective of differences among documents. These patterns will vary across individuals. Since objective techniques are less objective than they seem, and subjective responses of individuals are subject to variance, we had to find another way of combining understanding with representing the document in its own terms. This led us to think in terms of an institutionalised style of proceeding which would minimise variation in interpretation whilst accepting that interpretation is the central intellectual activity involved in information processing.

At this point the two stage approach to preparation of descriptions we had devised when working with PRECIS, and had subsequently retained for convenience, began to take on new meaning. Such an approach was needed in the PRECIS context so that subject experts could study our descriptions in relation to original documents. Since we could not expect them to read all the originals, we prepared fairly full summaries and coded required methodological information on a well piloted 'worksheet' (see further paper in this section) to serve as document substitutes. Descriptions were prepared on the basis of worksheets and compared with them. We continued to prepare worksheets, even when the original reason for them had gone, simply as a handy first step towards descriptions, allowing discussion within the research team of appropriate content and organisation of descriptions. Preparation of the worksheet was viewed in a way equivalent to the information scientist's statement of subject in that, though more extended, we did not believe that any other skill was involved than the technique of summary.

A major change in our document processing techniques occurred when we realised that this stage of processing could not be taken for granted in this way, and must be regarded not as a simple point of departure where information processing begins, but as a complex process requiring study in its own right. We first became aware of some of the factors involved in the course of batch preparing of descriptions on the basis of worksheets after a lapse of time in which memory of documents, summarised early in the sample, had blurred. Some worksheets, taken at face value, appeared to contain non-sequiturs of various kinds (eg variables which theory suggested no reason for including, arguments which apparently could not follow from premises). Whilst in a few cases errors had been made, and the worksheet had served a useful purpose in drawing attention to them, more commonly the non-sequitur was apparent rather than real, and additional information in the original made the reasoning clear. However, without this information, users would probably dismiss a given study as ill-conceived or as not making sense.

Further insights emerged when we tried to think out guidelines to specify what it is necessary to include in a worksheet to overcome this problem. We realised that particular pieces of information, or categories of information, have no intrinsic importance. They receive their importance from their relation to the theme and purpose of a particular investigation, and understanding of the overall content of a document, conversely, is validated in terms of internal consistency of the particular elements it comprises. To take a very simple example, race as a variable along with age and social class in a sociometric study of friendship choice in a coeducational school might seem of no more importance than any other variable, unless there is reason to suppose that the study was intended as a contribution to understanding of race relations. If this were the case, document analysis which led to the production of a description in the form of a simple concatenation of concepts would miss the point of the study, and either mislead the user or 'lose' the document for him. If there were doubt as to the author's central purpose, study of the way he handled his variables in analysing his data would confirm or disconfirm the point. The issues are often much more complex and sometimes, for lack of evidence, must be left to the user to determine. It is of interest that we are not unique in our thinking in this respect. The Human Relations Area File system at Yale, for instance, accepts nothing less than a reproduction of the total document as context required by the user.

Our experience indicated to us that document analysis should be handled as a formal stage in information processing, distinct from description writing, and with particular procedures designed to take account of the factors involved. This meant that the preparation of descriptions came to be viewed in a different light. With regard to document analysis and the preparation of worksheets, the way in which different pieces of information give one another meaning within the whole context of ideas is extremely complex. As this became more obvious, so we found that our approach tended to focus on elements that could be excluded, leaving the basic ideas untouched, rather than the reverse. The paper on preparation of worksheets explains this point further, although we have only just begun to explore this approach in any detail. The preparation of descriptions, in line with this thinking, could no longer be conceived of in terms of analysis or of imposing structure, appropriate or otherwise. This would be to disrupt the structure we have been at pains to preserve in the prior process. Descriptions must thus be regarded as a further reduction in which the 'next least relevant information' is excluded. This process of successive reduction has the effect of providing a means of solving the problem of seeing the wood amongst the trees to which information scientists have drawn attention. A certain amount of streamlining of words and structure appears to be acceptable for ease of comparision of items by the user, and a range of formal patterns has begun to emerge as optional though not mandatory for use, singly or in combination, in preparing descriptions. We would stress, however, that there should be no more than streamlining.

Subsequent work on our indexing scheme will undoubtedly influence our procedures in that, by viewing our documents against particular contexts of use, definitions of least and less important will vary around the basic sets of ideas in the documents. We have too a considerable way to go to make explicit the intellectual processes which operate in information processing. Even with the background of understanding shared by subject experts engaged in the task, the processes need to be made as explicit as possible if they are to be effectively employed for the particular task in hand. Whether it will be possible to explicate them in such a way as to sensitise non-subject experts to the ways in which subject experts view and think about their documents, and to acquire understanding by 'doing' such thinking, without a guided programme of formal study of the subject in question, is not clear at the moment. We can see no reason in principle why this should not be as feasible as internalising the style of thinking assumed by any existing information system.

III 5. (ii) Symbolic logic as an aid to preparation of descriptions, 1971.

1. We have experimented with the techniques of symbolic logic as an aid to subject analysis and statement. Our general conclusion is that this is a most valuable tool in extracting the overall structure of ideas in an article and in clarifying the relationships amongst concepts. It appears to us to be less helpful as a notation in which to express subjects, except perhaps at an additional intermediate stage prior to string writing. It has however suggested an alternative technique for presenting the subject in natural language which may be more helpful than our present one.

2. The main problems we have found in expressing our subjects in terms of propositional calculus are:

(a) Our subjects for the most part strictly contain only premises. This means for instance that our 'given...then...' pattern is strictly 'if...and if...'.

(b) Even where it seems valid (c.f. Elder on problem sheet) to express the subject as having, in logic, a quasi-conclusion, this results in a repetition of concepts which is undesirable in indexing:

e.g. $((\mathrm{ORC} \equiv \mathrm{RCRO}) \cdot ((\mathrm{PLGA} \cdot \mathrm{PLNS}) \supset \mathrm{RCRO}) \supset ((\mathrm{PLGA} \cdot \mathrm{PLNS}) \supset \mathrm{ORC})$*
where ORC = Orientation to racial change
PLGA = Perceived likelihood of goal attainment
PLNS = Perceived likelihood of need satisfaction
RCRO = Responses to racial change strategies of civil rights organisations

(c) This example shows also that we are using material implication ($\supset$) in a way which lacks logical rigour, rather in the sense of 'has implications for'.

(d) Consequently, in the statement of even relatively simple subjects, alternative and seemingly equally defensible expressions can be suggested, as an example from amongst those we have studied on the problem sheet illustrates:

Dean

(1) $((\mathrm{F} \supset \mathrm{ER}) \supset \mathrm{SR})$
(2) $(\mathrm{F} \supset (\mathrm{ER} \supset \mathrm{SR}))$
(3) $(\mathrm{F} \supset (\mathrm{ER} \vee \mathrm{SR}))$
or just possibly
(4) $(\mathrm{F} \supset \mathrm{ER}) \equiv (\mathrm{F} \supset \mathrm{SR})$

Where F = Fisher
ER = Ed. reform (compulsory secondary ed. etc.)
SR = Social reconstruction

(e) For indexing purposes, particularly in more complex subjects, it is necessary to consider the relationships of concepts within 'concept clusters' such as 'perception of role of language', which are often prepositional phrases. For this purpose it would be necessary to turn to predicate calculus and this would (unduly we feel) introduce further complexity.

---

* Symbols used

Material implication	$\supset$
Conjunction	$\cdot$
Disjunction: weak	v
: strong	V
Material equivalence	$\equiv$

3. These considerations seem to represent considerable drawbacks and led us to wonder whether the rough workings which we made as a preliminary to our attempts to express subjects in the notation of symbolic logic would not in themselves be sufficient to clarify the basic structure of ideas, allowing the logic of the subject rather than formal logical relationships to dictate the order of constituent phrases, and leaving within-phrase order to linguistic usage. Perhaps the most complex subject on our problem paper is Parsons and Platt, which was analysed as follows:

A Phases of social change in social structure and individual values

(taken in conjunction with) — Association of A with B

B Interrelation between personality differentiation and phases of individual life course

(is seen as an explanation for) — Association of (A with B) with (C with D)

C Growth of higher education involving extension of socialisation process

(since this is linked with) — Association of C with D

D Change in social structure of ed. and in individuals' values in the phase of post-adolescence

This kind of presentation and the type of analysis it involves is similar to, though less rigorous than, a formal application of logic but may perhaps serve our purpose adequately enough (incidentally also obviating the need for reference to a key to symbols adopted for individual concepts or phrases).

III 5. (iii) Experiment with automatic production of descriptions, March, 1972.

The basic aim of this paper is to compare the index entry produced by a two stage system (involving the reading of the full text by a subject specialist) and an entry produced by various short-cut techniques (involving a greater proportion of simply clerical time).

The adequacy of the subject descriptions as produced by the research team is taken as a given (representing required output) and various methods of reducing the initial reading time involved will be compared with this in terms of whether or not this description could have been produced from reading extracts, of various lengths, of the full article.

The methods to be compared with the reading of the full text are reading the following, or combinations of the following, which could be extracted from documents by a clerical assistant:

1. The title.
2. 1st section/last section (if document divided in this way).
3. 1st paragraph/last paragraph.
4. 2nd paragraph/penultimate paragraph.
5. 1st sentence of every paragraph.
6. Last sentence of every paragraph.
7. All sentences with clues to structure of the argument, eg firstly, lastly, in conclusion. There are three points .... Aims, objectives .... All the evidence seems to point to .... Main, central, focal, important.

The sample of documents is small (5) but was selected in order to provide a spread of document types with which our techniques must cope. These include one document expounding a sociological theory of social structure, an empirical study of students, two theses relating to the teacher's role and student unrest and an empirical study used as a basis for the development of theory. For each document the bibliographical details and the subject description prepared by the research team will be followed by comments on each method of extraction.

One particular point worth bearing in mind when comparing the various methods is the basis on which any consideration of accuracy should be made. Broadly it appears that an extract could be deemed adequate if it could provide the concepts stated in our descriptions. However, this alone would make it impossible to link the various concepts together as is required by the description, and therefore it is not only necessary to consider the adequacy of the extracts in terms of whether concepts are present but whether sufficient information is given as to the relationships between the concepts.

A. British Journal of Sociology, 1970, 21(1), 86-94.
Kemeny, P.J. Dualism in secondary technical education.
Investigation of relationship amongst social background, occupational placement and attitudes to education of secondary technical school boys; implications for assessment of effectiveness of policy for secondary technical education. Source of data - sample of 151 boys (at age 20) educated at two secondary technical schools in England.

Cont'd....

1. The title

This would have given very little information except the level/type of education under consideration. Dualism, without further qualification, gives no further specification to the problem.

2. 1st section/last section

This is a little more helpful than title. It would have revealed the notions of expected/actual functions of secondary technical education and the implications for policy. However, it would not have revealed that an empirical investigation was carried out and, obviously, gives no indication of the variables or sample involved.

3. 1st paragraph/last paragraph

This is similar in outcome to 2.

4. 2nd paragraph/penultimate paragraph

Here the kind of study to be carried out is hinted at in the 2nd paragraph where the variables are set out, in broader terms than they appear in the description. However, one could still not state with certainty that an empirical study is to be carried out, specify the variables accurately or specify the sample.

5. 1st sentence of each paragraph

This method, because it retains the basic structure of the article, is superior to the earlier ones. The background considerations are outlined, the notion of a survey to investigate these and the variables involved are specified. However other variables eg reasons for taking 13+ appear to be given equal weight and it is difficult to determine relative importance. Sample information is incomplete, giving age of boys and number of schools but not number of boys. Implications are specified.

6. Last sentence of each paragraph

This method appears totally inadequate in this case as no consistent thread appears to be maintained, explanations often being given without what is to be explained. However, in combination with the previous method it would have been sufficiently adequate to provide the variables of the study, how they were related, the level of education involved, the implications and full sample data. The additional information comes from two sentences only and this again raises the question of relative importance of 'peripheral' data.

7. Key statements

One of these statements gives an indication that an empirical study is to be carried out and indicates both the variables studied and the level of education under consideration. However, no indication of the context of the study (ie why is it being carried out), its implications or the sample is given. Words stressing importance of particular points, 'ie 'important' government reports, were selected as well as those giving a clue to the main focus of the paper.

Cont'd....

B. The Sociological Review, 1969, 17(3), 415-437.
Jessop, R.D. Exchange and power in structural analysis.

Outline of a conceptual framework (synthesis of conflict, consensus and exchange models) for analysis of social structure, distinguishing between centre and periphery of social system, analysing power (control over four types of means corresponding to political, social, cultural and military substrata) and exchange relations, and considering types of structural differentiation (between substrata in terms of means and control, and within substrata through division of labour); illustration of application of model.

1. The title

This would give a completely misleading impression of the document for it suggests that it is concerned with the value of two concepts in the study of structural analysis rather than the presentation of a synthetic framework, involving these concepts, for such analysis.

2. 1st section/last section

This would give the basic argument that a synthesis of theories is to be outlined for the purpose of analysing the social structure. Similarly the idea of application of the model to problems emerges. What is lost here is richness of description for no details are given of distinctions made and thus it would be difficult to distinguish this from other studies proposing a similar synthesis.

3. 1st paragraph/last paragraph

This is similar in outcome to 2.

4. 2nd paragraph/penultimate paragraph

In this case the author begins on his specific argument in the second paragraph ie outlining one of the models he is to synthesise and in the penultimate paragraph is still dealing with specifics. Hence on the basis of these paragraphs a totally incorrect idea of the document is given.

5. First sentence of each paragraph

This provides a better solution than 2 in that it gives the possibility of identifying all the major distinctions which are specified in the description. However a number of other variables are introduced eg exploitation, alienation whose relative importance is not made clear by the extracts. The illustration of the model is lost by this method.

6. Last sentence of each paragraph

This extract presents such a disjointed view of the document as to make it impossible to come up with an overall structure of the argument. However all the variables in the description occur in the abstract. The problem seems to be one of how to relate them together and how to ascertain relative importance of variables.

7. Key statements

Here, neither the models to be synthesised nor the purpose of the synthesis are revealed and although all distinguishing variables were mentioned their relationships were lost.

Cont'd....

C. New Society, 1969, 372, 769-771
Frye, Northrop. Anarchism and the universities.
Thesis concerning anarchist nature of contemporary radicalism (seen as central element in American culture, and as associated with decline in ideological sense resulting from religious crisis); implications for measures to deal with student unrest.

1. The title

This includes two of the concepts in the study but gives no indication of the relationship between them. In fact if one guessed at the relationship ie anarchism in the universities the result would be a misrepresentation of the document.

2. 1st section/last section

Not applicable to this document.

3. 1st paragraph/last paragraph

The first paragraph gives only an historical background to the problem and the last has a 'preaching' function. Hence, neither could be useful for preparing a description.

4. 2nd paragraph/penultimate paragraph

In the second paragraph the historical background to the problem is further considered and in the penultimate one a specific point in the argument is outlined. Hence, neither would be useful for writing a description.

5. First sentence of each paragraph

All the right 'words' appear to be included somewhere in these extracts but the relationships between them are impossible to formulate. The problem of peripheral data is highlighted here due to the presence of a number of illustrations, eg concerning literature, which tend to blur the overall structure of the argument and make it difficult to judge relative importance.

6. Last sentence of each paragraph

Here the problem mentioned in 5 is even more acute and many of the sentences have a prophetic tone which means the main threads of the argument are lost.

7. Key statements

Only two statements were extracted from the document by this method. Both were selected because they contained key words (main and great) but this emphasis was simply to stress a particular point in the argument and not to state an important focus of the study.

Cont'd....

D. Harvard Educational Review, 1969, 39(2).
Cuban, Larry. Teacher and community.
Thesis that effectiveness of teachers in inner city community schools depends on a professionalism which involves not merely instruction but also playing a part in curriculum materials development and close involvement in community activities; model for training programme proposed and compared with traditional programmes; implications in terms of role conflict, and costs and benefits.

1. The title

This picks on the two key elements in the study but does not show how they are related. There is an obvious loss of detail between title and description and even with the title a number of possible interpretations of document content could be made eg relations between home and school, teacher's participation in community activities etc.

2. 1st section/last section

These two sections are long and provide most of the data required to prepare the description. However the notion that a model for teacher training is actually presented is only approached indirectly in that benefits of the model are given but its nature and purpose are not specified. This is possibly due, to some extent, to the fact that the model is presented in diagrammatic form and thus not fully expanded in the text.

3. 1st paragraph/last paragraph

Neither of these paragraphs contains any information relevant to the content of the document for the first simply provides a vivid introduction and the last attempts to justify the inadequacy of the article.

4. 2nd paragraph/penultimate paragraph

The second paragraph gives some social background to the problem to be considered and the penultimate one presents some conclusions related only to a small part of the article - the efficiency of the programme proposed.

5. First sentence of each paragraph

This approach retains something of the structure of the argument and would provide all the concepts required to write the description. However it would be no easy job to link the concepts together in an acceptable way. The suggestion that a model of teacher education is proposed is present, but the fact that this is compared with traditional models is lost.

6. Last sentence of each paragraph

Here the loss of structure of the argument makes it almost impossible to make sense of the statements. The tripartite nature of the teacher's role, the fact that a training model is proposed and the implications are not specified. Many of the statements provide prophesies rather than 'facts'.

7. Key statements

All but one of these statements included some important aspect of the argument and on the basis of these the notions of the tripartite teacher role and costs and benefits could have been found. However the particular stress on community activities and the training model would have been lost.

Cont'd....

E. Sociology, 1969, 3(1).
Bernstein, Basil and Henderson, Dorothy. Social class differences in the relevance of language to socialization.
Presentation of model of social learning which links differential emphasis on use of language, different areas of orientation (basic skills/ interpersonal relations) and different forms of social relations within the social structure, reflecting different implicit theories of learning (self regulating/didactic) which affect child's concept of self and of role relationships. Leading from investigation of social class differences in perception of relevance of language in familial socialisation, distinguishing between interpersonal socialisation and socialisation into basic skills. Implications for sources of discontinuity for the working class child between home and school. Source of data - sample of 100 mothers in England.

1. The title

This presents an adequate statement of the empirical part of the paper although losing some of the richness found in the above description. It gives no indication of theoretical considerations, sample used or implications suggested.

2. 1st section/last section

The first section (introduction) gives a good account of the empirical work to be reported, the fact that a sociological explanation follows from this work and some indication of sample (the numbers were not given). The theoretical part of the paper would have lost in richness, ie specified as 'a sociological explanation of social learning in terms of the mediation of the linguistic process in socialization', if this extract had been used as the basis of description.

3. 1st paragraph/last paragraph

These give no indication of the content of the document except at a very broad level ie ... language acquisition, the relationship of language and cognition and the social antecedents and regulative consequences of forms of language use.

4. 2nd paragraph and penultimate paragraph

These extracts are far superior to 3 for providing information on the content of the document but although the empirical content is fairly adequately covered, as is the sample, no indication of the theoretical notions presented are to be found.

5. First sentence of each paragraph

The empirical part of the paper is well covered but although the presentation of a model is suggested no details are given and hence considerable richness would be lost. This problem may have been accentuated by the diagrammatic representation of the model which was not fully explained in the text. Sample information is present but is misleading as it states a total sample of 311 which was in fact the population from which the actual sample was drawn. The extracts are heavily overweighted with methodological details because these sections were stated in short paragraphs of one or two sentences. This resulted in almost complete reporting of the analysis of variance techniques employed and the methodological criticisms of the schedule used.

Cont'd....

6. Last sentence of each paragraph

These sentences were so disjointed as to give no indication of structure to the argument and although the variables from the empirical study are mentioned their relationships would be difficult to formulate. The theoretical part of the paper does not come out at all.

7. Key statements

One of the sentences here gives the three hypotheses to be tested in the empirical part of the study and these could have been used to write an adequate description of what was done. None of the other sentences gives additional important details and no clue to the theoretical background, implications or sample is present.

Summary of preferred extraction types for each type of document.

1. For item A, representing an empirical study, a combination of methods 5 and 6 (1st and last sentences of each paragraph) appears to be the only feasible solution. All methods taking extracts from the beginning and end of the document gave no indication of the empirical nature of the study, the title gave only level of education and the key statements ignored background considerations while identifying the empirical problem.

2. For item B, representing a pure sociological theory study, method 5 seems most appropriate providing data on the main aim of the study and the major variables. The title proved misleading and inaccurate whereas methods 2 or 3 (first/last paragraph/section) were accurate but limited to broad considerations.

3. For item C, representing a theorising article, only methods 5 and 6 have any value if a description is to be written for only these retain the main line of argument. The moralising content of 6 leaves 5 the only real solution.

4. For item D, representing a thesis, method 3 (first/last sections) was valuable acting as a genuine introduction and conclusion to the article. However this method omitted important elements of the study. Only method 5 covered all the elements.

5. For item E, representing an empirical study leading to a theoretical proposition method 3 or method 5 would have provided an adequate description of the document but in no case would it have been possible to give any details of the theory proposed. Even the title, in this case, gave a fair understanding of the empirical content of the document.

Hence the kind of study seems to make little difference to the appropriateness of any particular method. As might be expected those items containing introductions and conclusions (D and E) gave better results for the methods selecting extracts from the beginning and end of documents than those items without these sections. Overall, method 5 appeared most appropriate for all types of study.

Cont'd....

Conclusions

None of these methods of mechanically extracting sections from total articles either alone, or in combination, would have made it possible to have written any one of our descriptions. The inadequacies of output would have ranged from distortion and omission to loss of richness and detail. Even in those extracts where all major variables were specified, in most cases it would have been impossible to relate these together accurately.

The best compromise, in all cases, was the method involving first sentences of each paragraph. This superiority is probably a product of standard English which encourages the specification of the key idea to be considered at the beginning of each paragraph, the length of the total extract so produced and the fact of selecting sentences throughout the argument making it almost impossible for any major idea not to be given, at least, a minor mention. The latter two, of course, were present in the method of selecting out the last sentence of each paragraph but this proved far less helpful. The big problem with this method concerns relative importance of ideas. This exhibits itself in a number of ways. In some cases illustrative data are given equal weight to main arguments, methodological data may be overstressed because they are stated in short paragraphs and a number of variables may be introduced whose relative importance is difficult to determine without further context.

The present experiment shows that if a mechanical method is to be used to cut down reading time it will need to be specified in much more detailed terms. A mechanical system taking no account of meaning appears to be completely useless except as a method of extracting key words for relationships between variables cannot be formulated. The only approach that took any account of meaning was the last one but this proved a complete failure perhaps because instruction laid too much stress on key words (such as important) rather than centrality to the argument. Perhaps one possibility would be to see what key statements would be picked out by a 'layman' responding to meaning cues as well as syntactic cues.

Similarly such meaning cues would need to be chosen in such a way as to select extracts throughout the document so that no major idea, even if this was only specified in one section of the article, would be completely missed, while allowing for peripheral or illustrative data to be left out.

A further requirement would be the need to specifically state methodological information that was deemed important. In no case did sufficient methodological information emerge and even when it appeared to, as in Bernstein, the information was incorrect.

Another big problem arose where illustrative diagrams or schema were included in a study. These were normally not fully explained in the text and thus vital information was lost. Perhaps all charts etc. would need to be supplied for any short cut to indexing.

Cont'd....

The big question arises, then, concerning whether such an approach, based on a search for meaning, albeit a lay search, together with the specification of additional details, would really save time. The disjoint output of any straight extraction procedure makes it necessary to use considerable thought to link sections together and may in fact take as long as the reading and summarizing of the article by the 'expert' who is to write the description. Mechanical extraction too, tends to blind one to any major themes which may have been missed by the 'layman' and thus additional time must be spent in validating procedures.

Note

As specified earlier the present procedure for preparing descriptions involves a two stage process. Firstly, the item is summarised, using authors sentences or part sentences, on a worksheet. Further specific pieces of information eg country of study, sample, methodologies are added to the worksheet. This is then used as a basis for writing descriptions in a structured form ie theory, empirical content, implications and sample. The worksheet summaries, in all the present cases, are shorter than the extracts produced by methods 5 or 6 though, in general, longer than the other methods of mechanical extraction. They also give some notion of relative importance of ideas. Although the following analysis suggests that total sentences are reported, this is not always the case and additional links may be built in to summaries, within square brackets, if the points do not appear to follow from each other.

For illustrative purposes two items are selected for a detailed consideration, one representing a thesis (item D) and one representing a theoretical study (item B).

Item D comprises 4 sections containing a total of 62 paragraphs spanning 19 pages. Sentences or part sentences were extracted for the purposes of the worksheet as follows:-

Section 1 comprising 5 pages and 13 paragraphs

Page	Paragraph	Sentence
2	3	1st and last
	4	1st
	5	3rd
3	6	1st and 4th
	7	2nd and last
	8	1st and 2nd
4	9	1st
	10	1st and 3rd
	11	1st
5	13	5th and last

Cont'd....

Section 2 comprising 9 pages and 25 paragraphs

Page	Paragraph	Sentence
1	1	Title and 1st
2	3	1st
3	4	1st
	5	1st
	6	In toto (5 sentences)
4	7	2nd
6	12	1st
	13	2nd and 3rd

The remaining 12 paragraphs were by way of illustration ie accounts of interns experiences on teaching courses and therefore account for no extracts.

Section 3 comprising 5 pages and 15 paragraphs

Page	Paragraph	Sentence
1	1	1st
2	3	1st
	5	1st and 2nd
	6	2nd and 3rd
3	7	1st and 3rd
4	14	Last sentence

Section 4 comprising 3 pages and 9 paragraphs

page	Paragraph	Sentence
1	1	1st

Summary

One thing that appears obvious from this outline is that no consistent pattern appears to be present, although there is a predominance of first sentences of paragraphs (18 out of 38 sentences) included and more entries came from the introductory section than from the other sections. What also emerges is that one extract from any one paragraph is as common as more than one, suggesting that in many cases more than one extract is required to present a point of the argument. Similarly extracts tend to come from restricted 'blocks' of the document again suggesting that a number of extracts are required to illustrate points. One large part of the article provides no extracts because it was purely illustrative giving experiences of interns on training course proposed.

Item B comprised 11 sections containing a total of 59 paragraphs and spanning 16 pages.

Section 1 comprising 2 paragraphs

Page	Paragraph	Sentence
1	1	5th, 6th and last
	2	In toto (1 sentence)

Cont'd....

Page	Paragraph	Sentence
Section 2 comprising 5 paragraphs		
	No extracts	
Section 3 comprising 2 paragraphs		
1	1	1st
Section 4 comprising 4 paragraphs		
1	2	1st and 3rd
Section 5 comprising 5 paragraphs		
	No extracts	
Section 6 comprising 7 paragraphs		
1	2	1st and 3rd
	3	1st
	4	1st
	5	1st
2	7	1st
Section 7 & 8 comprising 20 paragraphs		
	No extracts	
Section 9 (a summary) comprising 6 paragraphs		
1	1	2nd
2	3	2nd & 4th (duplicate S.4) 5th (duplicate S.5), 6th.
	4	2nd & 6th
	5	1st
Section 10 comprising 7 paragraphs		
1	1	2nd & last
Section 11 (conclusion) comprising one paragraph		
1	1	6th

Summary

All the points relating to the previous item are also applicable here. It also shows the essential need for flexibility in any system of extraction, for no mechanical system could respond to the presence of a summary in the middle of an article (as occurs here), while also taking account of additional information which this summary does not provide.

DW/RAC
29.3.72.

III 5. (iv) The treatment of material relevant to the sociological analysis of education by existing information services.

It is extremely difficult to comment, in a general way, on the manner in which existing services handle material relevant to the sociology of education. Existing services vary immensely in their aims and in the assumptions concerning the nature of users and information which underlie the formulation of such aims. Document description can mean anything from presenting a list of key words (only one or all of which may be present at any given point where reference is made to the document) to providing a complex pattern of related concepts. A service may aim simply to provide users with a mass of possibly relevant material, leaving them to follow up and select from the total list those which are actually relevant, or it may aim to take over this selection function for the user.

Any particular service if looked at from the point of view of its own aims is likely to be adequate. Nevertheless when indexing services are viewed against general criteria, such as ease of access of relevant documents by users, in almost all cases defects are apparent. It frequently appears that document descriptions give the user an unhelpful or ambiguous view of the meaning of the document, although they are rarely incorrect on points of fact. A brief scan of the output of some existing services yielded a number of specific examples of misrepresentation. Some of these are classified below.

1. Description incorrect in fact

Very occasionally a description will be completely wrong. This is the case with the description for Oxtoby, R. "Educational and vocational objectives of polytechnic students", British Education Index, 1971, 7(6). One of the headings chosen was 'vocational guidance'. The article dealt in fact with vocational courses in universities and polytechnics and considered whether polytechnics should or should not be purely vocational institutions. The use of this heading could merely be a mistake, of the kind likely to be encountered in preparing any index, arising perhaps from a misunderstanding of the document; however, it could point to a somewhat more serious problem: if no heading existed relating to the vocational content of courses, the choice of vocational guidance might indicate that the nearest thing in terms of words was chosen. Assumptions underlying what constitutes a valid or discrete heading may influence the way a particular concept comes to be expressed. This is certainly the cause of a number of somewhat less severe errors that have been found. A number of examples from the British Technology Index will illustrate this point. Where achievement motivation is mentioned in a subject it is automatically translated into 'achievement' and where conceptions of teacher's role are studied, this is translated into 'Teaching. Aims: Opinion'. Both of these examples show how the degree of fragmentation employed in creating valid concepts for use as headings places severe limitations on the expression of meaning. Such problems are by no means limited to the particular index mentioned.

2. Matching

The assumption that the user locates documents by a process of matching appears to underlie most existing services. The user is expected to translate his problem into concepts which he then searches for in the index. If he does not find the concept he requires he is directed by means of cross-references to other concepts which are considered to be synonymous with the one which he originally selected. The user may be inconvenienced by a synonym list which is not exhaustive or by a system which invariably offers a cross reference when he attempts to enter the index via particular concepts. If the individual knows a document to have an index entry he will normally find it eventually although even this may be a time-consuming business. Take for example Smithers, Alan. "Students experience of thick sandwich courses", British Education Index, 1971, 7(5). A valid concept analysis of this document might be 'Perceptions of education', 'University courses', (involving) 'Industrial experience'. The actual entry for this document was 'Industry: training within industry'. There were no other educational headings. Without a complete familiarity with the list of headings employed this description of a document would have been difficult to locate even when one knew that it existed. In the situation where one is generally browsing through relevant headings much work of importance might be missed due to a concept analysis of one's problem, which is inadequate in relation to the terms used in the index.

One specific assumption underlying the notion of matching is that ideas can be validly broken down into individual concepts and located by combining such concepts. This tends to result in references to those documents whose descriptions contain the same words being located together in the index. In many cases this is not valid. Take for instance the complex concept of 'perceptions of the role of language in socialisation'. This might well be broken down into simple concepts such as language, socialisation and perhaps attitudes. The document itself would have only passing relevance to any of these headings and would not belong with other documents under such headings. If the user searching for such a document were familiar with the kinds of headings available, he might well be able to piece together his topic, but not without a considerable amount of cross-checking. The only other alternative would be to read through all possibly relevant headings.

A related problem is deciding what constitutes a relevant heading. This is aggravated where there are no agreed definitions of the meanings of the various headings. The sociologist is presented with a particularly difficult problem for his technical terms tend also to have lay meanings. This often means he must attempt to forget his subject background if he is to use existing services in a beneficial way. One prime example is the combined lay and technical use of the term 'social structure' in the classified section of the British National Bibliography. The sociologist is forced to plough through a mass of material on such topics as Women's Institutes and horse-riding clubs in order to locate the few documents dealing with the sociological usages of the term.

The relevance of a document for the individual might also be a product of the particular context of the concept being dealt with and not simply the concepts themselves. In the case of Salt, John. "Isaac Ironside, 1808-1870: the motivation of a radical educationalist", British Education Index, 1971, 7(5), the heading 'England and Wales: History of Education: 19th Century' gives no indication that this document might have relevance to the sociologist or social historian. The implication of the heading plus the title is that the document is a biographical study of one educationalist. In fact the document has much to say about the social movements of the period and the ways in which social and educational movements interacted.

Thus the assumption of matching appears to lead existing services to separate related material, bring together unrelated material and in general make location of documents difficult, particularly for the subject expert.

3. Hierarchical structure

Most existing services are based on the assumption that any given concept belongs to a limited number of hierarchies of terms. Cross references reflect this hierarchy by directing users from general to more specific headings. This assumption also influences an indexer's decision concerning what headings to select to specify a particular concept in a document, if that concept is not itself a valid heading, with the result that inappropriate generalisations are often made. Take for instance Christie, T. and Griffen, A. "The examination achievements of highly selective schools", British Education Index, 1971, 7(5). Here the notion of selective contexts for education is rounded up to 'Secondary Education: Selection' ('Attainment' is the other heading chosen.) This appears to result from some sort of implicit hierarchy. In this particular case a more appropriate summary term would have been related to the social aspects of the problem: the key factor is the effect on performance of working in the context of other people of high or mixed abilities. A similar error is made in the case of Stradling, Robert and Zurick, Elia. "Political and non-political ideals of English primary and secondary school children", British Education Index, 1971, 7(5). The headings selected were 'Politics and education' and 'Attitudes'. Again the choice of politics and education appears to be a product of an implicit hierarchy concerning the words politics or political .... To speak of politics and education implies some sort of structural relationship between institutions. Political in the context of this document refers to the extent to which children of different ages took political figures such as the Prime Minister or the Queen as role models in preference to pop stars or teachers. Hardly political in the structural sense.

To return to the issue of cross references. Existing services do not allow for the possibility of horizontal cross referencing. Thus items on different hierarchies are not generally linked by a cross reference. 'Related' references do help the user to move across hierarchies but are generally inadequate for they are limited by the idea that given concepts are only relevant in the context of a small number of hierarchies. In a multi-disciplinary field, like the sociology of education, absence of such a procedure might be particularly unhelpful. Imagine for instance a hypothetical example of a study of the effect of family background and ability on achievement in America and India. Such a topic might well find itself classified as achievement (or attainment in British Education Index), social factors, measured ability, America, India, cross-cultural studies. The last three headings represent a hierarchy which would imply that America and India were taken as countries to be compared. To the sociologist this article may be particularly relevant for understanding the process by which an individual's status is arrived at in societies based on sponsored and contest mobility. To cater for his requirements a horizontal cross reference from achievement to mobility and from, for instance, America (as a society based on contest mobility) to sponsored and contest mobility would be essential.

4. Nature of headings selected

In many cases existing services appear to choose headings which have empirical referents and to select such headings in preference to other types of headings if there is a choice. This often leads to severe distortion of documents. A particular failing is where the focus of a document is on some phenomenon with no empirical referent but an example is used which can be described in such terms. In the case of Gross, Neal. "Implementing organizational innovations: a sociological analysis of planned educational change", British National Bibliography, Jan-April, 1972, the particular entry in the classified arrangement is 'Educational innovation; United States; Study examples: Primary schools; Research reports'. The sociological content of the ideas underlying this particular work is almost completely unspecified. In the PRECIS index there is a reference from 'innovation' to 'educational innovation' but no downward reference from social change which is the more likely concept to be scanned for this kind of material by the sociologist. This, in addition to the fact that the PRECIS string does not indicate a sociological content, is likely to result in the sociologist overlooking the item by either not finding it at all or viewing it as irrelevant. Similarly with Williams, G. "Are more dons worse dons?", British Education Index, 1971, 7(5), the heading selected was 'Universities and colleges: teachers'. In fact the article dealt with the relationship between expansion in higher education and teacher quality and the validity of using qualifications as a measure of such quality, both in fact with no obvious empirical referent. Other examples include the work of Freire, Paulo. "Cultural action for freedom", British National Bibliography, October 25th, 1972, where remedial education for illiterate adults is taken as the subject, without any reference to the more hazy, though nevertheless central notion that one prerequisite of democracy is that the individual should be free from the bondage of illiteracy. The sociologist concerned with this topic would probably overlook this book under this heading.

This appears, in part, to be due to indexers lacking expertise in particular subject areas. This explanation gains some support from looking at examples where no empirical referent really exists. Take the example of Goffman, Erving. "Interaction ritual: essays on face-to-face interaction", British National Bibliography, Jan-April, 1972. This can only be accessed via 'Social psychology; Public interpersonal relationships; Essays'. The middle concept appears to be an invented one aiming to get at the technical content of the document but creating a nonsensical entry point as far as subject experts are concerned.

Because phenomena with no empirical referents are not widely considered as the basis for headings, no attempt is made in traditional indexes to distinguish by authors perspective within a discipline. Take for instance the items edited by Young, Michael F.D. "Knowledge and control: new directions for the sociology of education" and by Hopper, Earl. "Readings in the theory of educational systems", both to be found under the same main heading, 'Social institutions', in the classified section of the British National Bibliography for Jan-April, 1972. These generated the headings 'Education; Sociological perspectives; Conference proceedings' and 'Education;. Sociological perspectives; Readings', respectively. In other words, they were seen by the indexing system as virtually identical. However in terms of particular perspective they are totally unrelated, representing two opposing threads of sociological theory, namely a sociology of knowledge and a sociology of input and output structures. These books would address totally different questions, on the basis of totally different assumptions but to the index and, thus, to the uninformed user, they are the same.

Thus assumptions about the objective nature of headings lead traditional indexes to ignore important distinctions between documents and to focus on often spurious similarities between them.

5. The nature of concepts in documents

Up to this point existing indexes have been considered from the point of view of their assumptions concerning the nature of concepts and user procedures in searching for them. Only passing reference has been made to the notions of complex concepts and no consideration has been given to the feature of some indexes which attempt to indicate the relative importance of concepts or how they are linked together. In the majority of cases existing indexes appear to ignore the possibility that a concept formed from a number of smaller constituent units might be essentially different from the sum of its component parts. Many sociological concepts do, however, involve larger units as illustrated by the example quoted previously of the 'perception of the role of language in socialisation'. Another aspect of sociological work is that subjects of documents often involve relationships between concepts rather than the concepts themselves. Take for instance the example of Armer, Michael and Youtz, Robert. "Formal education and individual modernity in an African society", Current Index to Journals in Education, Jan-June, 1971). The concepts used as headings included social change, changing attitudes, developing nations, but the focus of attention was not in any of these specific factors but on the extent to which individual modernity in the context of a developing nation is a function of, ie varies with, formal education. No existing services appear to cater for the possibility that the focus of a document might be the relationship between concepts rather than particular concepts themselves.

In view of such problems it would appear that indexes offering a variety of headings or attempting to indicate relationships amongst them would be more valuable to sociologists than those offering only one heading at any particular mention of the document.

Current Index to Journals in Education attempts to indicate the relative importance of concepts by presenting together two lists of relevant concepts, the most important being indicated by a * marking. Unfortunately the length of the list of important concepts is arbitrarily limited to a maximum of five and also the level at which such concepts are pitched appears to be determined by the broadest level of concept being found in the document. In the following example: Kandel, Denise and Lesser, Gerald S. "School, family and peer influences on educational plans of adolescents in the United States and Denmark", Current Index to Journals in Education, Jan-June, 1971, limitation on the number of important headings means that only broad headings are starred. These are cross cultural studies, socioeconomic influences, academic aspiration, adolescents and cultural differences, whereas the important concepts for this particular study - peer relationship, parent influence and school role are unstarred.

In the case of: Astin, Alexander W. "The methodology of research on college impact, part one", Current Index to Journals in Education, Jan-June, 1971, the starred concepts say everything that is important in this document at a fairly specific level. The additional headings are irrelevant in the context of the overall focus of this particular document for they involve the people and properties being studied whereas the focus is on methodological issues.

Thus, even in a system attempting to indicate the relative importance of concepts, ambiguity of meaning is not overcome. Such systems, still leave the user very much on his own when it comes to deciding on the relevance of any given document for his own particular needs.

DW/RAC/19.12.72.

III 5. (v) Working notes on preparation of worksheets, Easter 1973

The purpose of a worksheet, apart from ensuring that at least one person has studied a given document in detail, is to give a balanced and condensed representation of the document. A copy of the pro forma used is appended to this paper, together with instructions for completing it.

In principle nothing is irrelevant, if only in that it contextualises the main content, and so nothing can be excluded. In practice, we want to be able to take a snap-shot view, and we must therefore identify and exclude the least important statements or passages. There is no absolute way in which any element is or is not important. Clearly this depends on the overall structure and content of the document.

There are certain common-sense principles such as one applies to note-taking or precis-ing in any situation. One may, for instance, include main points to a level which gives a clear idea of the overall complexity of ideas, but exclude subsidiary points below this level, examples which do not advance an argument or add new factors to a definition, or other detail. Or one may omit purely introductory or concludory material, ie material which does not constitute a starting point to which an author returns, or specify an approach which throws light on the work as a whole, or serve some other specific purpose. One may exclude data included to quantify, if a general statement covers it. And so on.

There are certain preferences which may be specified. For instance, if a general point is made, elaborated and then repeated, 'general point + elaboration' is preferable to 'elaboration + general point' from the point of view of ease of comprehension. Where alternative statements of variables are given, the most complete one as related to the operational definitions is essential, but a more generalised statement is also helpful. Where the author makes a point to which he returns at some later time, and this is essential to the argument, it is helpful to repeat it.

However, beyond such obvious and practical points, it is neither desirable nor possible to give cover-all rules. To do so would be to lose what is distinctive about many documents. In particular, the material we are handling is very various in character. At the same time institutionalisation of the process of worksheet preparation beyond that already built into the worksheet pro forma may be required so as to avoid idiosyncracy. A decision making guide such as exists for selection would probably be so complex as to be unusable if it were to be wholly adequate for the purpose, although something of a very openended nature is a possibility for the future.

For the moment we can only illustrate the kinds of principles by which we work. A first reading of a document prior to consideration of what will be included in/excluded from the worksheet is a prerequisite. This will elicit any general statements as to the author's central concern (eg proposing a theory and testing it, stating a position and justifying it, focusing on a situation and explaining it, identifying a practical problem and seeking a solution for it, selecting a theme and providing information relevant to it). There is also commonly some indication of the audience addressed or of what the author has attempted to do by way of improvement on or difference from related work (eg question fundamental assumptions, synthesis, greater rigour, draw attention to factors previously overlooked, provide relevant data). It is largely this sort of appreciation of the document which determines what one 'sees' in a document and includes in a worksheet - it sensitises the reader to cues indicating eg how ideas are related to one another, what the main points are and what are subsidiary points, whether the same point is being made in different words or whether there is an important difference.

So, to take an obvious example, an author concerned with alternative definitions of wastage may well be contrasting two statements differing rather subtly in meaning, and neither can be excluded. Another writer discussing wastage as a problem requiring action may use two very similar statements but will, by virtue of the focus of his interest, attach no importance to the difference, or even to the two statements which could both be excluded if subsumed under a more general statement, unless they contextualise some later point.

In general, the kind of thinking we are seeking to institutionalise is one which refers not to general rules, nor even requires decisions to be made in the context of predetermined questions (although such sets of questions may emerge in due course), but rather provides for questions to be formulated in relation to and on the basis of, the individual document, decisions being related systematically to such questions, and based not on an individual's beliefs about what an author is attempting to say but on internal evidence (eg if earlier he has said X, he must at a later stage be doing Y and not Z).

The foregoing applies to section C of the worksheet. Little guidance can be given with regard to the methodological sections following. These depend upon familiarity with the research techniques employed in the field.

Some revision of the worksheet is planned. This is likely to involve a somewhat different approach to the content sections, and to affect detail only in the methodological ones. This is being deferred pending further development of our indexing scheme, since we may wish to extend and possibly reorganise the worksheet in the light of such development.

Appendix A to Paper III 5. (v)

OSTI/SEA PROJECT WORKSHEET

A BIBLIOGRAPHICAL DETAILS (SEA format): a Author(s) and Title:subtitle

b Journal articles – Journal, Year, Vol.(issue), first–last pages.Refs.

or c Books, reports etc – Place of publication, Publisher, Date. No.pages,Refs.

For office use

B CONTEXTUAL INFORMATION, IF GIVEN, ABOUT:

a Author

b Document

C CONTENT OF WORK (indicated by key statements from document)

a Level 0

i Sociological or 'quasi-sociological' concepts and their interrelationships, together with any definitions necessary to understanding work

ii Type of conceptual framework (see code)..........................

b Level 1

'Real world' (education or other setting) problem – general statement, indicating author's standpoint if he has one

c Level 2

i Particular circumstances to which consideration of problem is limited

Temporal..

Geographical: Country...

Type of setting...

Educational level...

Social unit: Level of analysis................................

Nature of unit................................

ii Particular aspects considered other than those indicated above

OSTI/SEA PROJECT WORKSHEET (cont)

ABBREVIATED BIBLIOGRAPHICAL DETAILS Books - Brief title, Year;
Journal articles - Name of journal, Year, Pages.

For office use

D NATURE OF WORK (see code):

a Main approach (if there is one) b Other approaches (instead or as well)

....

If coding in D above was 2 or 3, complete rest of sheet; if not skip to next sheet

E OPERATIONAL DEFINITIONS OF VARIABLES (specify)

F TYPE OF DATA (see code)..

If either of these apply, note only any difference in treatment of data from that in original study; if not complete sheet in full

N.B. The following items refer to handling of data as data, not to interpretation (see codes for each section)

G TYPE OF CASE STRUCTURE

a Size and composition of sample (specify)..........................

b Structure of sample...

H CONTROL OF VARIABLES

a Temporal..

b Situational...

c Specify controlled variables......................................

....

J DATA COLLECTION

a General methodology...

b Technique...

c Instrument..

K DATA ANALYSIS

a Procedure of analysis...

b Test of significance..

OSTI/SEA Bibliography and Index Project - Guide to filling in description sheets and to coding schemes

The description sheets are intended to contain all the necessary information on which selection decisions will be based and validated, and from which index entries will be prepared, without the need for further reference to the originals.

MAIN WORKSHEET

This provides general information about a document under the following headings:

A BIBLIOGRAPHICAL DETAILS Presentation in SEA format (excepting only affiliation, see below), which includes all the details uniquely to identify a given work. One item extra to SEA requirements is an indication of number of references cited in bibliography (or as an alternative a statement of whether footnote references are given).

B CONTEXTUAL INFORMATION

a About author Eg academic or professional background, institutional background or status. Only such information as is given in the article Can be ignored if such information is not given or if writer is well known.

b About document Including reasons for which work was conducted (eg underdocumented area) or circumstances in which reported (eg thesis, paper read at conference).

C CONTENT OF WORK Indicated in the main by statements extracted from the original and edited as little as possible to avoid any distortion. Divided into what, for want of a better term, we have called levels.

a Level 0 has to do with, by analogy with Chomsky, the 'deep structure' of ideas, ie concepts represented in their most generalised form. These will normally be sociological concepts, with the occasional addition of 'quasi-sociological' concepts drawn from socio-economic, social psychological etc thinking.

i We are interested in:

The concepts used,

How they are used (eg family as social unit, family as institution),

The relationships seen to exist amongst these concepts.

ii The type of conceptual framework which orients the writer's thinking will also be indicated, coded as follows:

Sociological

Basis of conceptualisation (its organising principle)

1 Conflict theory/Marxism
2 Structural functionalism/(Social) systems theory (ie analysis of society in terms of 'system')
3 Symbolic interaction
4 Other (specify)
5 Not clear

Mode of conceptualisation

1 Neo-positivist
2 Phenomenological/Intuitionist
3 Other (specify)
4 Not clear

Other disciplines

If other discipline framework is used specify eg social psychological.

Also state particular type of theoretical framework if indicated eg Socioeconomic. Institutional (c.f. Veblen).

b Level 1 At this level the statement of content consists of what Riley[1] calls empirical indicants of concepts or manifestations of problems or situations in which 'real' phenomena are expected to behave in the ways which the researcher's theory predicts. However our bibliography will also include a number of studies in which 'real' problems are investigated or discussed for their own sake and without reference to the generalised concepts with which the sociologist works, though these may be implicit. In these cases level 0 will be left blank, although later inferences may be made. In these cases also any general statement the writer makes about his problem will be indicated, eg as in prescriptive writing 'Need for...', 'Desirability of...'.

c Level 2 will be completed where the problem under consideration is not treated comprehensively. Topics may be delimited in various ways - see coding below. It is to be expected that something that forms the main topic of one document may be treated as an aspect of a more general topic in another document.

i For convenience various ways in which topics may commonly be delimited are listed and coded as follows:

Temporal

1 Retrospective (specify period) NB ignore historical 'introduction' and occasional historical 'examples'
2 Current (ie 1944 Act onwards)
3 Comparison between periods (specify)
4 Not clear

Geographical

Country (specify)

Type of setting

1 Urban
2 Suburban
3 Rural
4 Combination
5 Other (specify)
6 District or town if nature of district is not stated
7 Not clear

Educational establishment or level (specify, eg primary school, CAT)

Social unit (specify, eg school, neighbourhood)

NB Whilst the categories under each of these sections are intended to be mutually exclusive, a single document may deal with a number of things and require multiple coding.

ii Further kinds of delimitation are, in sociological works, the ways in which the data has been analysed eg by age, by sex, etc, and in the case of non-sociological work, by factors which are seen to form part or to have a bearing on a problem. (Specify).

---

[1] Riley, M. W. Sociological Research. New York, Harcourt, Brace and World, 2 vols., 1963.

D NATURE OF WORK REPORTED IN DOCUMENT There is provision to indicate whether work is of a single kind, eg theoretical exposition, whether there are other additional approaches, eg details of empirical study, or whether there are several approaches none of which is uppermost.

Coding for 'Nature of work'

1 Theoretical

2 Empirical - hypothesis testing or exploration which generates hypotheses (often impossible to distinguish between these)

3 Empirical - fact finding

4 Factual (bringing together known facts or statistics)

5 Interpretative ('more' than 'factual' but 'less' than 'theoretical')

6 Descriptive (giving more unified picture than 'factual', may be impressionistic)

7 Prescriptive (arguing one side of case, making 'ought' statements)

8 Literature review (commenting on state of knowledge)

9 Critique (commenting on individual's work)

X Bibliographical (offering no assessment of work, purely descriptive, possibly merely listing)

Y Other (specify)

THE FOLLOWING SECTIONS (E - K) APPLY ONLY TO EMPIRICAL WORK

For the most part the coding schemes, which owe much to Riley's[1] paradigms, are self-explanatory.

(They are preceded by ABBREVIATED BIBLIOGRAPHICAL DETAILS as a safeguard in case related sheets become detached from one another.)

E OPERATIONAL 'DEFINITIONS' OF VARIABLES Ie how variables are operationalised, eg social class/father's occupation. (Specify.)

F TYPE OF DATA

Previously **reported** data

1 From empirical work - raw data

2 From documentary sources other than raw data of experimental work

If either of these apply, note only any difference in treatment of data from that in original study; if not complete sheet in full

3 Data collected for the purpose

4 Other

5 Not clear

G TYPE OF CASE STRUCTURE

a Size and composition of sample (specify). Its relation to the population represented should be indicated if this is clear. If not it will be taken that the social unit indicates the frame of reference.

b Structure of sample

1 Random

2 Stratified/Cluster

3 Quota

4 Convenience

5 Other (specify)

6 Not clear

H CONTROL OF VARIABLES

a Temporal

1 Single time

2 Through time (specify)

3 Not clear

---

[1] Riley, M.W. Op.cit.

b Situational

1 Systematic (specify)

2 Partial control (specify)

3 No control

4 Not clear

Specify variables controlled or indicate, if not clear, that this is the case

J DATA COLLECTION

a General methodology

1 Interaction analysis

2 Operations research

3 Systems analysis

4 Action research

5 Simulation

6 Sociometric study

7 Other (specify)

8 None special

(These are general approaches which may influence the kind of data collected, by contrast with Ka below, which at most will concern structure of data.)

b Technique

1 Participant observation

2 Non-participant observation

3 Questioning

4 Other (specify)

5 Not clear

c Instrument (Specify if named instrument is used, or else indicate general type of instrument eg questionnaire, interview, diary.)

K DATA ANALYSIS

a Procedures of analysis (Specify eg cluster analysis, factor analysis, regression analysis. Include here different types of rating or scaling, specifying if possible as Likert, Guttman etc. These are included under data analysis although the encoding of data is done in effect by respondent.)

b Tests of significance

1 Chi square

2 Other (specify)

3 Not clear

Appendix C to
Paper III 5. (v)

OSTI/SEA Bibliography and Index Project - Summary of coding schemes

A BIBLIOGRAPHICAL DETAILS
B CONTEXTUAL INFORMATION
  (a) About author. (b) About document
C CONTENT OF WORK
  (a) Level 0
    i) We are interested in the concepts used, how they are used (e.g. family as an institution) and the relationships seen to exist among them.
    ii) The type of conceptual framework which orientates the writers thinking will also be indicated, coded as follows:
    Sociological
    Basis of conceptualisation (its organising principle)
    1. Conflict theory/Marxism.
    2. Structural functionalism/(Social) systems theory i.e. analysis of society in terms of system.
    3. Symbolic interaction.
    4. Other (specify).
    5. Not clear.
    Mode of conceptualisation
    1. Neo-positivist.
    2. Phenomenological/Instructionist.
    3. Other (specify).
    4. Not clear.
    Other disciplines
    If other discipline framework is used specify e.g. socio-economic, socio-legal etc. and state particular theoretical approach if any is indicated e.g. socio-economic institutional (c.f. Veblen).
  (b) Level 1 : Problem stated in empirical terms.
  (c) Level 2 (i)
    Particular circumstance to which consideration of problem is limited.
    Temporal
    1. Retrospective (specify period).
    2. Current (i.e. 1944 Act onwards).
    3. Comparison between periods (specify).
    4. Not clear.
    Geographical
    Country (specify)
    Type of setting
    1. Urban.
    2. Suburban.
    3. Rural.
    4. Combination.
    5. Other (specify).
    6. District or town if nature of district is not stated.
    7. Not clear.
    Educational establishment
    Specify e.g. primary school, C.A.T. etc.
    Social Unit
    Specify e.g. school, neighbourhood etc.
    Level 2 (ii): Further kinds of delimitation of problem.

D NATURE OF WORK REPORTED IN DOCUMENT (coding)
  1. Theoretical.
  2. Empirical - hypothesis testing or exploration which generates hypotheses.
  3. Empirical - fact finding.
  4. Factual.
  5. Interpretative.
  6. Descriptive.
  7. Prescriptive.
  8. Literature review.
  9. Critique.
  X. Bibliographical.
  Y. Other (specify).

THE FOLLOWING SECTIONS (E - K) APPLY ONLY TO EMPIRICAL WORK

E OPERATIONAL 'DEFINITIONS' OF VARIABLES
  Specify how each variable is operationalised
F TYPE OF DATA
  Previously collected data
  1. From empirical work - raw data.
  2. From documentary sources other than raw data of experimental work.

If either of these apply, note only any difference in treatment of data from that in original study; if not complete sheet in full

  3. Data collected for the purpose.
  4. Other.
  5. Not clear.
G TYPE OF CASE STRUCTURE
  (a) Size and composition of sample (specify)
  (b) Structure of sample
    1. Random.
    2. Stratified/Cluster.
    3. Quota.
    4. Convenience.
    5. Other (specify).
    6. Not clear.
H CONTROL OF VARIABLES
  (a) Temporal
    1. Single time.
    2. Through time (specify).
    3. Not clear.
  (b) Situational
    1. Systematic (specify).
    2. Partial control (specify).
    3. No control.
    4. Not clear.
  (c) Specify variables controlled or indicate, if not clear, that this is the case
J DATA COLLECTION
  (a) General methodology
    1. Interaction analysis.
    2. Operations research.
    3. Systems approach.
    4. Action research.
    5. Simulation.
    6. Sociometric study.
    7. Other (specify).
    8. None special.
  (b) Technique
    1. Participant observation.
    2. Non-participant observation.
    3. Questioning.
    4. Other (specify).
    5. Not clear.
  (c) Instrument (Specify if named instrument is used, or else indicate general type of instrument e.g. questionnaire).

K DATA ANALYSIS
  (a) Procedures of analysis (specify).
  (b) Tests of significance
    1. Chi square.
    2. Other (specify).
    3. Not clear.

III 5. (vi) Preparation of descriptions: progress report, December 1972

Note: We use the term 'description' to refer to a condensed representation of a document, which will stand in lieu of the document in our system. This term has been adopted in preference to 'index entry' or 'abstract' which tend to be associated with particular modes of condensation. We have regarded mode of condensation as a matter for investigation, not necessarily one of selection from existing options.

Early experimentation

Our initial attempts at description writing were conditioned by the fact that we were attempting to work within the PRECIS indexing system developed by the British National Bibliography. Subsequently, an effort was made to develop a system sharing only certain general characteristics with PRECIS, but nevertheless capable of meshing with it.

Both these indexing approaches proved to involve unacceptable constraints upon the form and content of the document representation. The main kind of constraint which is relevant in this context was the requirement that the document representation be framed so as to link it meaningfully with each subject heading assigned to it (in the case of PRECIS the representation is used for the automatic generation of subject headings by computer manipulation).

Since this time, our theoretical work has led us to a conception of a literature searching system alternative to that which appears to underlie PRECIS and like systems. We have argued that the imposition of structure confers meaning, and that meaning determination is not properly a task which a system can undertake for its users, because it cannot enter into their minds. However, if we are to facilitate the literature searching process we shall necessarily impose structure. In practice, therefore, the effective system will be one in which structure (1) is kept to a level which does not intervene seriously in the search process, and (2) is sensitive to the range of user perspectives rather than rationalising perspectives away or ignoring them.

Later experiment

The decision to develop an approach to accord with the particular nature of work in our field did not point immediately to appropriate procedures. Rather than build in preconceptions ab initio, we spent a period in which we tailor made the description to the individual item, so that future practice could be later decided on the basis of insight and experience of the problems presented by our material. Descriptions were circulated to the authors of the items handled to check on possible distortion (no serious criticisms were levelled at our efforts).

With several hundred items behind us, we are now in a position to review this work and discuss alternative procedures we might establish. With several hundred more items part processed and awaiting the preparation of final descriptions, we have studied the issue with considerations of the operational (as contrasted with the 'laboratory') situation in mind. The review process has also taken us back to our theoretical framework as well as causing us to look forward to the operational situation.

Thus two kinds of criteria have been weighed in the review process: intellectual appropriateness, and practical viability. The latter kind of criterion would not of itself justify any set of procedures; it constitutes rather a major factor in evaluating theoretical alternatives. We have to distinguish the system and the user standpoint in considering practical viability; again system viability has no value in itself.

Theoretical considerations

Our conception of a literature searching system rests on the assumption of user interaction with documents and not only, as in conventional systems, on an objective matching of documents and search requests independent of the user's intellectual processes. *A priori* formulation of search requirements cannot be assumed in our field. We hope later to identify a range of user strategies and build them into the system, and this would influence description writing practice. But even then the requirement of minimum intervention holds, lest we distort user/document interaction by destroying cues, in the process of condensation, which may be relevant to the user's search process. We think now in terms of a user *devising* a strategy; we hope eventually to enable him to *select* from a range of strategies which are meaningful to him.

In reviewing our procedures in the light of these ideas, it became clear that patterns (see appendix) had built up over time in description writing but, because not made explicit, were not altogether consistently adopted and not amenable to validation as such. This placed us in the dilemma that formalisation of our procedures to achieve consistency might introduce a rigidity which would be a source of distortion. On the other hand, if left implicit the patterns represented no less a risk of distortion, and if eliminated these patterns were liable to be replaced by others, quite probably varying across description writers.

Consideration of these alternative sources of distortion led us to realise that we had no basis on which to assess the seriousness of alternative sources of bias. The main options appear to be as follows (each assumes that the terms in which an author states his concepts or variables are inviolable, but represents a different position regarding intellectually appropriate overall structuring of the description and/or the terms in which relationships amongst concepts are stated):

(1) Structure as made explicit by authors.

Advantages. Misrepresentation in one sense at least is ruled out and the document is handled as unique.

Disadvantages. One is nevertheless selecting, and possibly 'scissors and pasting' from the text so that, for instance, important elements might be omitted or minor ones elevated in importance.

(2) Formalisation.

Advantages. Institutionalisation of bias, provided this is done in a way meaningful to users, means that bias is a factor which users can 'manage' in interaction with documents.

Disadvantages. From the system side such institutionalisation may lead to abuse in that procedures may become so internalised as to replace sensitivity to the individual document. Safeguards such as regular review of procedures may minimise this effect.

Note: This applies both at a macro level (the major structural units in a description, such as presentation of theory, empirical investigation as guided by stated theory) and at a micro level (eg relationships amongst variables posited by a theory or investigated empirically), as well as in the relating of macro to micro (eg the arguing of a principle as a basis for recommendations as to practical policy). In principle, it is possible to think in terms of different levels of formalisation, but it is difficult to justify partial-formalisation of any kind (if formalisation is acceptable, it can apply throughout, and if it is not, then equally this applies at all levels). In practice, too, such different levels tend to be interdependent.

(3) Subject expert's interpretation

Advantages. The situation in which an individual gives free expression to his perception of what the author is essentially saying is less tedious than practising the craft of constructing a description which the other methods require him to perfect; it may also be more personally satisfying.

Disadvantages. The license taken introduces personal bias which, at the level of the corpus of documents, tends to produce a 'scrambling' effect. Also, even if users are able partially to unscramble, they may be led to infer unintended differences in meaning.

Note: This approach cannot be avoided where authors are unclear or formalisation inappropriate or inadequate, but cannot be defended except as a last resort, even if fairly frequently brought into play, when processing is done in house. Personal interpretation becomes the only resort, but still only a resort, when practising subject experts contribute copy as in the Sociology of Education Abstracts situation. In these circumstances, we find that users regard it as an acceptable resort, ie when known experts are the source of material, because bias can be predicted. Users thus take over the task of quality control which would normally be a system function.

Practical considerations

The in house / in field organisational arrangement is possibly the major practical constraint on policy. Questions of the relative time involved in alternative options are not decisive, since differences are likely to even out once a description writer has been fully inducted. A concern to make provision for quality control, and the consideration of eventual meshing in with other information services, both tend to tip the balance towards formalisation. A service intended for the use of subject experts rather than information experts might well however find formalisation a deterrent to use. Sensitivity to the total range of perspectives which users collectively might bring to bear on the material in the system is a question apart. Formalisation could be geared to prior classification in terms of perspectives and is not ruled out in principle. However, this begins to put a construction on 'minimum intervention' which may depend on invalid psychological assumptions. There are the further questions of differences in the extent (in degree and/or kind) to which user search strategies deviate from their normal course (ie with complete documents to respond to) in relation to different modes of intervention, together with the extent to which deviation of one sort or another matters subjectively to users or more objectively to the practice of the discipline.

Empirical questions

In the situation where the structuring of descriptions is restricted by the variety of ideas employed by authors, problems of inter-'indexer' consistency are particularly important. It may be that individuals respond in unique ways to the problem of what to write when faced with a summarisation of a document. One factor which might be significant in this process is cognitive style or the manner in which an individual is predisposed to process information in particular tasks. Some support for the existence of such tendencies has emerged from introspective questioning concerning the way each member of the research team carries out the process of writing a description from a document summary. It was noted that one member preferred to take a holistic view - first identifying the overall structure of the argument in the summary, then extracting specific elements which seemed to contribute to such a structure, while a second took a more atomistic view - first recording all aspects that appeared important in the author's argument, then combining these to produce the total description.

Both these approaches involve an element of risk; the former, that the conception of the overall argument might cause important details to be ignored, the latter, that details of relatively low relevance which obscure the overall argument, might be included. There is no reason to assume that only these two possible styles exist, or even that any given indexer has only one such style. There may be many, and the difference in cognitive style from one indexer to another might lead to significant differences in the content of the descriptions they produce (in addition to the inevitable differences due to such factors as different use of language etc.). It would be unwise to assume either that such differences would or would not result in substantial differences from the meaning a given user would tend to attach to the original document. Such an issue can only be addressed by empirical investigation.

If differences in meaning were found to exist, some strategy for overcoming bias due to individual cognitive style might need to be adopted. In the particular case of the research team a possible solution might be for the pair of indexers, one of whom represented the holistic and one the atomistic style, to work on the same set of documents in order that discrepancies in descriptions could be discussed and descriptions modified. The position appears relatively clear cut in the case of our particular pair of styles, but many such styles may exist and lead to bias in descriptions. Without further knowledge of the range of possible cognitive styles and their specific consequences for descriptions, it is impossible to suggest strategies for overcoming bias.

A number of research questions suggest themselves in relation to differences in cognitive style. The nature of the phenomenon is perhaps the major question, for it is this that determines what form action may take. If cognitive style is shown to be a fixed property of personality, and therefore unchangeable, it is unlikely that any action, short of a wholly rigid formalisation of procedures (which has already proved to be impossible), can combat its effect at the description writing stage. Hence, any devised strategy would have to attack the problem after descriptions had been written. One possible, though hardly practical, strategy would be to have each author check alternative descriptions, by a number of indexers, of his own documents, and suggest which, if any, of the descriptions most accurately reflected what he was trying to say. If, on the other hand, cognitive style is not a fixed property of the personality, action may be determined by the particular aspects of experience which have led the indexer to his particular style or styles. Such styles may be determined, for instance, by subject background, (and hence familiarity with concepts and arguments employed in documents), or by length of experience of abstracting details from documents. If this were shown to be the case, strategies could be devised for eliminating the effects of style at the description writing stage, for example, by intensive training sessions in the manipulation of sociological ideas or in description writing. It should not, however, be automatically assumed that any differences which may exist in the cognitive style of individual indexers will necessarily lead to undesirable discrepancies in the descriptions which they produce. Much empirical testing may be necessary to indicate the conditions which could be said to have an undesirable effect in terms of users attaching meaning to descriptions. This, itself, would involve establishing criteria for determining which cognitive style or set of styles if 'best'. It may be decided that one particular style should be adopted for all descriptions, thus introducing consistent bias. On the other hand, it may be found that a particular style is most appropriate to a particular range of documents. Thus, the indexer might be required to 'switch' styles depending on the 'type' of document being considered. For example, one style may be appropriate for theoretical work and another for empirical work. Again, authors appear to be the only people who can determine the validity of any given style for handling their own documents.

Another range of empirical questions may be touched on in conclusion. These concern the meaning imputed to the document by the user, in contrast with the document as presented by the author. The principle of minimum intervention will not necessarily be satisfied simply by minimising distortion of what the author intended to convey. Subject experts may be expected to vary broadly, according to subject background factors (in particular, perspective), in the meaning they impute to a given document. If our descriptions are such as greatly to disturb this pattern of response, the principle of minimum intervention has not been satisfied.

There has not been time to carry out a prior investigation of user response defined in this way. This means that, even though we have been careful to present each document in its theoretical context, it would be a fortunate accident if we had managed to include in our descriptions all those elements to which users variously respond in the originals.

We have to ensure a broadly similar response to our descriptions as users would make in examining the original documents, if the system is not to 'intervene'. This means that descriptions should reflect the features of documents to which users respond. We need, therefore, to investigate the characteristics of user pattern of response to original documents. This will afford a measure of the adequacy of our descriptions from this point of view, and will point to specific weaknesses in our descriptions which we should remedy.

It might be that, collectively, users' perspectives more or less correspond with those of authors, and, in this case, the pattern of response will be relatively simple. On the other hand, if user perspectives differ from those of authors then the pattern of response is likely to be more complex.

It seems reasonable to suppose that users will often be familiar with the perspectives in the literature. This means that, even if their own perspectives do not correspond with those of authors, they will be able to translate their ideas into the language of authors' perspectives and in this way assess the relevance of other work for their own, or to translate others' work into their own language in order to assess relevance. However, we cannot assume that users' perspectives are always clearly formulated in advance (in fact we believe this to be the exception rather than the rule) and for this reason we need to study the process by which user perspectives interact with author perspectives during the course of a search. We shall not get insights into this process merely by studying the relation between users and authors perspectives as manifested in sets of document descriptions which form the product of searches.

Subject experts have pointed out two specific approaches in literature searching. There is the situation in which the academic has a research question, and he is looking for a language in which to think about it. There is also the situation in which the academic has a language which has produced helpful insights in his investigation, and he wishes to trace the genealogy of his ideas (this might be for the purpose of writing up his work or he might be concerned to consider the generalisability of his theory to substantive areas other than his own). There are undoubtedly others to be identified. We would expect these to be associated with different search strategies in use of our system. The choice of a search strategy employed might also possibly be mediated by other factors (eg subject background), which our descriptions might need to take into account. This would then put us in a position to consider the interplay between descriptions and other system elements in the search process.

15.12.72.

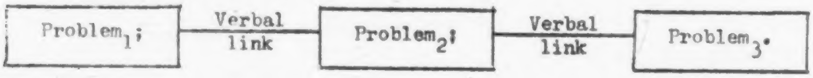
Appendix A to
Paper III 5. (vi)

Notes on formalisation

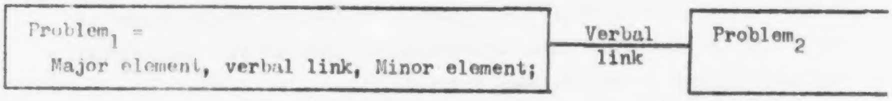
Overall structure of descriptions

It is helpful first to distinguish between two levels of structure. To avoid summoning up any preconceptions associated with terms such as 'subject' 'concept', we adopt instead less loaded research terms. The terms we use are:

Problem. A document may deal with more than one problem. This term may refer to various things - exposition of a theory, report of an empirical investigation carried out within a stated theoretical framework, implications of such work in terms of policy recommendations or new research questions raised. A new problem is initiated when there is a break in the intellectual structure and when an author 'changes gear' to move, for instance, from exposition to an empirical test of a proposed theory. Occasionally the problems dealt with within a document may be disjunctive, but more commonly there is an intellectual link between one and the next. Thus overall we can diagram a description as:



Element. Each problem may have major and minor elements. Major elements refer to the central research problem. Minor elements are those which do not stand on their own, but are considered necessary to an understanding of the nature of a problem as the author views it. Examples are important conceptual distinctions made, empirical areas on which he draws to exemplify a more general problem or argument. Minor elements are verbally linked to major elements, so that we can elaborate our diagram as follows:



Lists of recurring major and minor links follow. Additional links would be required for more unusual circumstances.

Major links:	Background to (context).
	Basis for (theoretical support).
	Leading to (chronological).
	Tested in (empirical support).

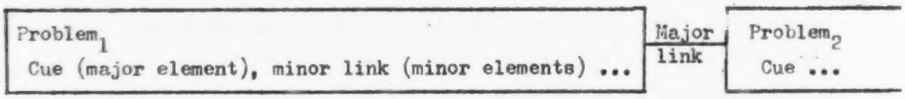
Minor links:	Drawing on data from	Commonly referring to situations.
	Extending to	
	From viewpoint of	
	In context of	
	In terms of	
	Stressing	
	Viewed as	
	With special reference to	
	Distinguishing between/amongst	Commonly referring to variables.
	Taking account of	
	With consideration of	
	By means of	Referring to purpose of study.

Another feature of each element is that problems may be presented in such a way as to cue them in with reference to nature of treatment. Examples of cues are 'theory', 'position', 'critique'. A list follows, and again the use of additional terms would be needed for exceptional circumstances.

Analysis (objective and rigorous examination of an intellectual position or social situation not associated with any particular individual(s)).
Comparison.
Critique (objective and rigorous examination of ideas of particular individual(s)).
Description ('factual' account of situation, setting etc, without comment on appropriateness etc. of its attributes).
Discussion (objective but not closely reasoned treatment).
Evaluation.
Implications for / that (providing they are followed through).
Investigation (empirical investigation - in special cases nature of study may be bracketed on eg Investigation (projective techniques)).
Outline (summary or preliminary account).
Overview (comprehensive account of state of thinking or knowledge at a point in time).
Position (case of thesis which is stated and supported by reasoned argument).
Presentation (exposition of case, theory, model etc.).
Recommendation (prescription, advice or, more tentatively, suggestions).

Review (refers to survey of the literature).
Statement (reference to or assertion of something taken for granted, eg statement of notion that).
Theory (may be either established or proposed theory).

Added into the diagram:



The sequence in which problems are stated most usefully perhaps follows that of the document. Both the number of minor elements qualifying the major element in a problem, and the number of problems comprising a description, are normally best limited to a maximum of three or four, having regard to the reader's powers of assimilation.

As a rule problems may be kept discrete within a description. However in a few instances (eg where a theory is coextensive with a supporting empirical investigation, and stated at more or less the same order of generality), two problems may be 'collapsed' into one. This has the advantage of shortening the description, particularly where there tends to be repetition of words. However, the disadvantages of greater length are often outweighed by the consideration of not conveying a misleading impression of the work. Collapsing is therefore determined by the question - will there be a loss of ideas, not - will repetition of words be eliminated.

Internal structure of problems

The major element is the basic unit of our descriptions, rather than the individual concept (cf the research design of a study v the individual variables under investigation)

Such elements may be associated with either an empirical treatment (ie there is a sample of subjects or units on whom data has been collected) or non-empirical. There is additionally a range of problems whose object is to present a particular position on some issue (eg that something is or ought to be so). Such work contrasts with both the former types of work in which propositions are expressed by reference to formal variables rather than a point of view concerning some 'variable' or 'variables'. The question of sample/not sample is of little relevance here. For example contrast the following:

Empirical

Investigation of emergence of social class differences in intellectual performance of black children ....

Non-empirical - formal proposition

Theory that high industrialisation/urbanisation is associated with low conformity to significant others ....

Non-empirical - point of view

Position that widespread image of contemporary youth as affluent, unconventional and swinging is inappropriate ....

We have found that, although there are broad differences across these different types of problems, it is possible to identify some broad categories of problems in terms of the research approach employed which are meaningful in all three contexts. The categories identified so far may be presented in the form of a typology, together with some linguistic variants associated with each. A full analysis has yet to be completed. Such categories can undoubtedly validly account for a substantial proportion of our material. It is not envisaged that the typology be made exhaustive (ie including all patterns down to sets of one), nor that it be fixed (regular review should be a feature of the system). It represents simply a partial record of precedents recognised at a given point in time. Its possible use in on-going description writing is a matter for system ground rules.

Some elements of such a typology are as follows:

1. Entities
   (a) Entity studied as given state at given point in time - seems inappropriate to 'variablise'.
   (b) Entity studied in terms of its properties - may be openended or characterised in relation to predetermined properties or analytical dimensions.

2. Properties
   (a) Relations amongst X, Y and Z) Neutral as
   Relation (between X and Y) to relation
   (of X with Y and Z respectively)

(b) (Effect of X (on Y)
((in inhibiting etc Y)
(Consequences of X for Y)
(X as factor in Y)
Differences in X (according to Y)
(analysed by Y)
(Y_1 compared with Y_2)
Congruence (between X and Y)
(of X with Y and Z respectively)
} Expectation of direction of relation

(c) X as function of (Y (ie X varies as Y varies))
(relation between Y and Z)
Relative importance) of X and Y (in Z)
Strength of association) (with)
} Expectation of source of variance.

3. Models and approaches

(a) Alternative conceptions / explanations } (classes of variables/analytic concepts) for X.

(b) Conception of X as Y)
X viewed as Y)

(c) X for understanding of Y

(d) Synthesis of X and Y

Thus, in terms of 'patterns', we may represent a fairly typical document as follows:

Theory concerning / A viewed as B,/ distinguishing between B_1 and B_2;/
basis for / investigation of / relative importance of P and Q in R,/
with consideration of S;/ implications for Z.
=
Cue / 3b, / minor link ...; / major link / cue / 2C, / minor link; / cue / 1.

Ancillary information

Finally, two further types of information may be appended to a description.

First, a writer may often suggest that his work has a bearing on a field of research or policy other than his own. This may be indicated.

Secondly, subject experts we have consulted attach considerable importance to certain aspects of research methods (eg size of sample) in evaluating a piece of empirical research in relation to their particular concerns. Draft guidelines follow.

Proposed guidelines for specifying methodological details

Methodological details are introduced by the phrase 'Source of data' followed by a dash. They comprise two distinct parts, separated by a full stop, which are details of sample and details of method.

1. Sample details

These generally comprise three context specifications: the human context or people studied, the physical context or institutions studied and the geographical context, or country of study. Each specification may be followed by further information enclosed in brackets.

In general, both human and physical contexts contain the same kinds of information. The number of units are specified, followed by details which apply to the whole sample which the author considers to be important (eg negro, or single sex) and finally the nature of the units is specified, eg children or schools (this order may be reversed in order to maintain standard English, eg persons occupying important positions). A bracket may then follow giving additional detail which distinguishes subsets of the sample. A standard division in relation to the human context is age (ie the bracket may be of the form 'ages 2 - 6'). In some cases the number of units of each type may also be included in the bracket (ie 2 grammar and 3 secondary modern schools).

In general the geographical context simply involves stating the country or countries in which the data are collected.

The three parts are generally introduced by the words 'sample of' and linked by the words 'from' and 'in', thus giving the final form of description as:

Source of data - Sample of 45 male students (ages 18 - 21) from one liberal arts college in US.

Exceptions to these general rules are frequent and arise for a wide variety of reasons:-

(a) More than one sample. Where different samples are being compared, or where distinct samples have been collected for some other purpose, they are generally kept separate in the methodological details, are introduced by the words 'samples of', and the distinction maintained as far into the description as is appropriate. If two samples of people share the same physical and geographical contexts, the details might be specified as 'Samples of 146 negro and 150 white students from one university in US. On the other hand, in the case of a cross-cultural study the distinction may be maintained throughout (eg Samples of 404 Oriental children (around age 10) in Canada and 460 Mennonite children (under age 15) in US).

In the case of control groups, their existence is indicated by a bracket, following the statement of human context, containing the words 'with control groups'.

(b) Geographical selectivity. The country or countries selected for the study might have been chosen because of some special feature they possess, eg developing countries. In this case the geographical setting is specified in terms of such features, the actual country names following in a bracket, eg in four Latin American countries (Chile, Brazil, Argentina, Peru).

(c) General rules inappropriate to specific sample. On occasions an author will state that he is studying units which do not involve one of the three contexts. If this is the physical context it is generally a simple matter to connect the human and geographical contexts directly. One convention which has been adopted, in cases where the sample is selected from the total population of a country, with no intervening physical context, is the introduction of the details by the words 'National sample' rather than simply 'Sample'.

If, on the other hand, the units of interest to the author do not involve a human context, problems may arise particularly if data are collected from human subjects. In many cases it will be clear from the words the author uses which units he is interested in (eg if he is concerned with studying organisational climates it is unlikely that any human respondents will be any more than a means to an end). In such cases it seems more appropriate to specify the sample in the form ... of six colleges in US. Because of the various difficulties involved in deciding which is the key focus of attention, and to indicate to users that such a study might be of interest to them if they are concerned with perceptions of the phenomena, the following convention has been adopted: the statement of units is followed by a bracket containing a statement of the number and nature of sources of information, followed by the word 'respondents' (in a study where verbal data collected) or 'subjects' (in a study where non-verbal data collected); eg Sample of 16 organisations (540 employee respondents). If no human sources are specified the same action, as for physical context, is taken and that part omitted. Such decisions will, of course, be in some sense arbitrary for whether respondents are specified depends entirely on the author who may, or may not, be particularly concerned with methodological matters. In cases where the focus of attention is not made clear and all contexts are specified it seems reasonable to retain the general format.

Occasionally an author's data will come entirely from documentary sources and in this case the sample details are specified simply as Source of data - Documentary sources, followed by a bracket specifying the particular source if this seems appropriate, eg 'Who's Who' or 'Census data'.

A further particular case arises with case studies of a particular unit or units. In those articles, where the author specifically states he is doing such a study, the sample description is specified as 'Source of data - Case study of ...'.

2. Details of method

So far this space has been reserved simply for information concerning the temporal dimension of data collection. However it is available for the specification of other methodological details which may be thought important. In the case of the temporal dimension, 'Longitudinal study' has been the most common statement. This can be followed by a bracket specifying the length of time covered by the study but only when such a period appears to be unusual in some way in relation to the study (eg Source of data - Case study of one secondary modern school in England. Longitudinal study (over 11 year period)).

Specific conventions for specifying methodological information

1) All numbers above ten are written in figures, numbers ten and below are written out in words. This rule applies except in the case of dates and ages in brackets where these will be specified in figures throughout. (If the particular age is unqualified it is assumed it refers to years of age. Otherwise the age must be qualified with m. for months or d. for days.) All numbers over one thousand have a comma between the hundreds and thousands positions.

2) Full stops are not used with standard abbreviations, ie US.

3) Possessive apostrophes will not be included in dates (eg 1960s, not 1960's).

4) Where details are not fully supported by the article, although some evidence, eg authors affiliation in the case of country of source of data, exists in the document they will be placed in square brackets together with a question mark, ie in [? England].

5) In all cases words seen as appropriate by the author to the country of origin of data will be employed to specify that data. They will not be converted into English forms, eg grades 3 - 7, elementary and high schools etc.

Note: these guidelines have not as yet been fully implemented. Successive sets of material have been used to pilot and refine procedures and practices. Full implementation of these guidelines is pending on editing of our total sample. Our descriptions at the present time are consistent only within sets.

AMERICAN SOCIOLOGICAL REVIEW

1970, 35(2), 253-267.

Rhodes, A. Lewis and Nam, Charles B. The religious context of educational expectations.

Investigation of relation between religious identification (mother's religious identification and religious composition of school) and college plans, taking account of IQ, family socio-economic status and mother's educational attainment. Source of data - Sample of 6,186 white students (ages 14-19) in US.

DAEDALUS

1970, 99(3), 568-608.

Metzger, Walter P. The crisis of academic authority.

Consideration of validity of various explanations for present crisis of academic authority, with particular reference to sense of exploitation as basis of protest and to religious form of student movements; discussion of guidelines for reform in staff-student relations; illustrated by account of setting up of communes at Columbia University in 1968.

HARVARD EDUCATIONAL REVIEW

1969, 39(2), 301-309.

Crow, James F. Discussion: how much can we boost IQ and scholastic achievement? Genetic theories and influences: comments on the value of diversity.

Critique of article by Jensen (relating social class and racial differences in IQ to genetic determinants) in terms of inadequacy of statistical method used for interpreting genetic differences in humans (randomisation techniques as used in experiments on animals) and of inadequate consideration of environmental factors as these differentially affect races.

JOURNAL OF SOCIAL ISSUES

1969, 25(2), 137-146.

Sandford, Reichart. A greater space in which to breathe: what art and drama tell us about alienation.

Thesis that contemporary alienated youth are like artists and dramatists in their reactions against social patterns and artistic beliefs of our time in that both are searching for new meaning and, like artists, can provide a creative and dynamic, though sometimes destructive, force essential to our society.

SOCIOLOGY OF EDUCATION

1969, 42(4), 386-399.

Edgar, Donald E. and Warren, Richard L. Power and autonomy in teacher socialization.

Theory of occupational socialisation as a power process in which attitudes of members joining an organisation change towards those of significant others (people with ability to control organisational sanctions), the process being modified by relative personal resources, and by degree of affect between individual and significant other, affecting motivation to change. Basis for explanation of socialisation change in teachers. Tested in investigation of relation between occupational socialisation and change in new teachers' attitudes to teacher autonomy (direction v delegation in a variety of task areas) relative to those of their significant others, influence of resources (teaching experience and perceived status of teaching subject) and affect (direct questions), with consideration of actual autonomy, perceived legitimacy and satisfaction, and taking account of sex and school level. Source of data - Sample of 638 experienced and inexperienced teachers and 89 administrators in one school district in US. Longitudinal study.

TRENDS IN EDUCATION

1971, 24, 17-22.

Hopkinson, David. The school leaving age.

Discussion of issues associated with raising of school leaving age (public opinion concerning advantages, response to the challenge by educational authorities and teacher training establishments, areas in which development is needed - buildings and equipment, teacher deployment, curriculum, teaching method and classroom organisation, and pupil attitudes in terms of returns yielded by education offered).

III 5 (vii) Authors' comments on descriptions

Authors' comments, 1971

1. Replies were sought from each author on:

   (a) Accuracy of the description of his article(s).

   (b) Details of any literature search(es) recently carried out or projected.

2. The request for details was sent out in association with a recent survey of research being conducted by the OU.

3. Response rate = 55% = 41 authors (without follow-up). A sample of our American authors is now being contacted.

4. In general there seems to be no major criticism of our descriptions. With regard to literature searches, data are now available to support decisions on subject headings and subheadings.

5. The following sections summarise both sets of replies.

6. Analysis of responses concerning descriptions

Satisfactory	28
Minor error (generally omission)	12
Major error (generally commission)	6
Additional detail suggested	21

Minor errors - The majority of minor errors occurred in the source of data section, either how sample should be characterised, or whether data were collected for the purpose or existing data were reanalysed.

Major errors - Apart from a couple of genuine errors, these involve questions of interpretation. There is also the problem of the statement of intent which does not necessarily appear to have been realised.

Additional detail suggested - In a few cases the need was felt for definitional information, but the majority of suggestions relate to questions of methodology which are felt to be significant (before/after, method, instrument, mode of analysis, etc.).

Further comments - Two writers asked for separate sentences rather than single multi-clause descriptions. Two others suggested that minor aspects of a study be referred to, on the grounds that these would be of particular interest to the audience even though the whole study was somewhat marginal (this raises the question of balanced representation).

Implications for description writing - These responses suggest that the only major area of policy which appears to require reconsideration is that of the inclusion of methodological details.

7. Analysis of responses concerning literature searches

(The attached sheet indicates what was asked of respondents.)

A general point is that on the whole respondents seemed to find no problem in describing their searches in the suggested pattern, suggesting that our predictions as to the factors which concern them were reasonably accurate. Only six answered a question other than that we asked, describing the 'channels' they would consult rather than the dimensions of the problem which were crucial in relation to the work in hand. (This may be partly due to contamination from another project which has been particularly concerned with 'channels', ie whether reviews of research/bibliographies/abstracts, etc. are the main sources of references.) There was also an occasional tendency to respond in terms of expectations of the literature rather than demands of the research, but it was encouraging that there was little inhibition in the sense of paring down the problem to what could be traced through available indexes.

As described, the search problems all fall into patterns of enquiry recognised in studying methods of content analysis of documents, thus lending further support for our description writing techniques. Roughly classified, they include studies of relationships amongst variables (12), analytical studies of situations or processes (8), methodological/theoretical studies (3), broad studies of research areas (2).

Only one half of respondents suggested under what subject headings they would hope to find relevant material. The suggested headings tended to be broad rather than specific, supporting the policy we are favouring.

With regard to dimensions of search, and the way in which material might be helpfully subgrouped under main headings, whole/part (eg Organisation/Authority structure, Decision-making systems) may sometimes be helpful. Theoretical perspective, level of analysis, group studied, all recurred (we suggested the latter two possibilities in our blurb). It is clear from the statements of problems that our 'other variable', subjective/objective categorisations might well be useful. Additional requirements were Whole/Types, Empirical/Theoretical, Quantitative/Non-quantitative, Slant (eg non-mathematical), Context, Application (eg substantive area to which study is addressed), Method (eg Before/After, Longitudinal, Instrument used), together with the obvious geographical or chronological split. Some of these are clearly more relevant to particular kinds of research than the organisation of material in terms of content. Most of these features can be taken account of when the user scans descriptions even if they do not form the basis of subdivision.

There appears to be no obvious link between nature of problem and type of subdivision required. This would seem to support the suggestion that material should be sorted several ways under a given heading to allow for different concerns.

DETAILS OF LITERATURE SEARCHES REQUIRED

We are interested in the following kinds of information:

1. Statement of problem.
2. Any particular approach or slant.
3. Any further limitations on your requirements (ie some material may be relevant in terms of problem and slant but not helpful for the particular task in hand).
4. Under what broad subject heading(s) would you think it appropriate for relevant material to be located in a search tool (please do not feel constrained to answer in terms of the practice of existing indexes etc. in answering this).
5. In what way(s), relative to your particular requirements, would you like the material under each heading to be subdivided, eg. according to unit of analysis, group studied etc. (Again please do not feel constrained to answer in terms of any indexes you have used.)

The amount of detail we require is illustrated by the following examples:

A "Studies of differences amongst school teachers which approximate to the differences between 'locals' and 'cosmopolitans' found by other researchers. Information required in connection with study of local-cosmopolitan differences amongst postgraduate students. Special interest in differences between those who see themselves primarily as teachers of subject and those who see themselves as teachers of children."

B "Background data upon parents' attempts to help in their children's education in the pre-school years. Their efforts in subjects which are part of the infant school curriculum - writing, counting and especially reading. Actual practices carried out by parents rather than 'intervention programmes'. Interest not in the voluminous literature on child rearing practices, child development and the intellectual stimulus offered by the home, but rather in specific attempts to prepare for the infant school curriculum."

C "Education and rural-urban migration: a) in period after World War II; b) internal migrations (within U.K.) only; c) education as both a cause and a result of migration with comparison of educational characteristics of movers and non-movers; d) impact of migration on education in sending (rural) and receiving (urban) areas."

Authors' comments, 1973

1. Replies were sought from each author on:
   (a) Accuracy of the description of his article(s).
   (b) 'Traditions' in which he was working at the time of writing the particular article(s).

2. It was explained that the above information was required for further research into the appropriateness of alternative methods of subject indexing for sociologists of education, on the assumption that an effective subject index must be sensitive to the structure of thinking in the field as defined by members of it.

3. Samples of authors with both British and American/Canadian affiliations were contacted. Response rate for British authors = 85% = 39 authors. Response rate for American/Canadian authors = 74% = 32 authors. (Both without follow-up.)

4. Only a small proportion of authors offered criticisms of our descriptions which would necessitate alterations in them. With regard to traditions, authors provided valuable information concerning the ways in which they think about their work in relation to the work of others.

5. The following sections summarise both sets of replies.

6. Analysis of responses concerning descriptions

Satisfactory	30
Minor error (generally omission)	14
Major error (generally commission)	5
Additional detail suggested	19

N.B. Each unit represents one description classified in terms of the most serious kind of criticism offered.

Minor errors - the majority of minor errors concerned presentation of the data included in the description, rather than interpretation of the original document. Some authors suggested alternative forms of words or punctuation. In some cases such alterations would interfere with in-house rules concerning the presentation of information in descriptions, and it seems fair to assume that, in such cases, if authors understood the rationalefor using particular forms of words, they would not wish the description to be changed. In no case would a minor error affect a searcher's decision concerning whether to locate an original document.

Major errors - There were only five of these. Two represented genuine errors in that important aspects of the particular document were not made sufficiently explicit. This would mean that a searcher would not be able to decide whether to locate the original document on the basis of such aspects. This also applies in the case of the other three, which represent a particular problem associated with the type of journal in which they were located: two were to be found in Educational Research and one in New Society. In each case, an aspect of the document which the author felt was not of central importance, was highlighted at the expense of more central aspects. Interestingly enough, one author pointed out that this was probably a result of his 'trying to say too much in too short a space'.

Additional detail suggested - Replies classified in this way represent authors who generally approved of the description of their document, but felt that in some way it did not go far enough. In some cases, the suggestions were in opposition to policy decisions. For instance, as a general policy, findings are not included in descriptions, but 5 responses requested this. The majority of the remaining replies required further definition of variables which were included in the description.

Implications for description writing - These responses, unlike those from the 1971 sample of authors, laid less emphasis on the desire for additional methodological details although, as yet, description writing policy has not been amended in this respect. This may be a result of the particular sample of descriptions for which responses were sought on this occasion. With regard to the present sample of responses, a significant point to emerge is that documents in particular journals may represent particular problems for interpretation and indexing. In relation to both these points, further empirical investigation is required before any policy decisions can be taken.

7. Analysis of responses concerning traditions

The majority of authors responded to this deliberately open-ended question. Individual authors tended to contextualise their work in a number of different traditions. In addition to the multiplicity of ways in which individual authors viewed their work, there were variations between authors concerning the way in which 'tradition' should be interpreted. One of the questions in which we were interested when making this request to authors was whether or not there is consistency within the sociology of education concerning the ways in which bodies of work in the discipline are categorised. Although there were considerable variations in the words used by authors to describe their work, certain patterns of response did emerge and for convenience the particular 'traditions' stated may be grouped under the following general headings. (Other forms of grouping are possible, for instance research/policy issues addressed, which may be more appropriate for organising the set of headings in terms of what authors had in mind.)

(a) General research approaches

- comparative research strategy
- eclectic
- empirical, empiricism, English empiricism
- formalist
- humanistic tradition
- liberal, liberalism, liberal/egalitarian viewpoint
- policy related research, policy implications
- practical
- pragmatic, pragmatic approach
- substantive areas
- theoretical areas, theory creation

N.B. Practical, pragmatic and empiricist approaches may be appropriately grouped.

(b) Discipline affiliations and general approaches defined in discipline terms

- child psychiatry
- clinical psychological
- comparative sociology
- cultural anthropology, social anthropology, socio-anthropology
- economic anthropology
- empirical psychology, data-based psychology
- historical sociology
- interdisciplinary approach
- mainstream applied social science
- medical sociology
- political journalism
- political sociology
- psychological ramifications
- social psychology/approach/level
- sociolinguistics
- sociology of education
- urban sociology

N.B. It might be appropriate to group all sub-areas in a major discipline (eg anthropology) together.

(c) General sociological theories

- anti-functionalist view
- 'civic culture' approach
- classical sociological tradition
- conflict approach
- exchange theory
- functional orientation, functionalism
- grounded theory
- historical/functional
- logical positivism
- reciprocal theory
- social interactionism
- structural analysis, social structural interpretation
- structural functionalism, structural functional
- symbolic/phenomenological, symbolic interactionism, symbolic interaction, symbolic interaction framework
- Weberian

N.B. It may be appropriate to group structural and functional approaches.

(d) General 'other discipline' theories

- political and economic anarchist thought
- psychoanalytic, psychoanalytic thought

(e) Specific sociological theories and models

- bureaucratic theory
- concept of institutional culture
- deterministic theories: accident theory and cultural theory
- genetic
- normative convergence model
- organisation theory, organisational theory, axiomatic theory of organisations
- reference group theory
- relative deprivation
- resource theories
- role theory
- social power, social influence approach
- socialisation theory, socialisation model
- stratification theory, stratification (attainment process)
- subcultural model

N.B. Organisation theory may be appropriately combined with bureaucratic theory.

(f) Substantive area

- assimilation-imperialism
- attitudes, attitude change
- collective behaviour
- college student environments
- communication, communications explosion
- community building and conflict
- education, educational process, medical education
- expectations
- family, family roles, sociology of the family
- group decision
- group structure
- human development
- interorganisational relations
- pre-occupational groups
- prejudice
- professions and occupations, sociology of professions, economics of professions
- psycho-social evolution
- race
- sex roles and effect of education, sex education, sexology
- social change
- social conflict, conflict
- social control
- social stratification, social class, class formation and consciousness
- social structure and psychological functioning
- socialisation, early childhood socialisation, adolescent socialisation, political socialisation, student socialisation
- sociology of existentialism, sociology of the absurd
- sociology of institutions and organisations
- sociology of meaning
- sociology of science
- sociology of sociology
- student unrest
- vocational choice

(g) Methodology

aggregative data analysis
content analytic
factor analytic
model building (mathematical)
multivariate
psychometric
quantitative measurement
survey, survey research, social survey

III Selected papers

6. Principles of overall intellectual organisation

The papers in this section fall into two groups. They represent only a small proportion of the material (both discussion papers and sample indexes) produced by the team in the process of experimentation with alternative principles of intellectual organisation. Where possible we have selected, for this report, summaries of phases of work rather than reports of individual experiments. A few examples of our indexes and classification schemes only are included, simply for purposes of illustration. A full set of papers and experimental materials will be made available on request, subject to OSTI's permission, and that of our Working Party, from the files of whose meeting papers most of the items are drawn.

(A) Work employing a subject heading approach

The papers in the first group take their starting point, following our work with PRECIS, from the realisation that our attempts to structure our subject headings within a single framework, be it simple or complex, had the effect of grouping like with unlike. The contexts in which concepts are used differ, and it seemed that our subject headings should be sorted on this basis before they could be meaningfully related together. The central notion which guided our thinking at this time was that of multiple frameworks, and we reasoned that both structure amongst subject headings and substructure associated with given headings would differ from one framework to another.

The first paper in the group describes some early general thinking about the nature of our concepts. This thinking was elaborated in detailed work on material relating to a particular cluster of concepts (achievement and related notions), and that work is summarised in the second paper. At a more general level, differences in clusters of concepts, in terms of the way the concepts were used in documents, seemed to suggest that we should view our clusters as forming several different frameworks of subject headings. The following four papers describe experimentation with alternative ways of operationalising the notion of framework. The results were not satisfactory. We came to recognise the inappropriateness, for our material, of objectivist assumptions implicit in the notion of subject heading.

4

(B) Work employing the notion of dimensions of knowledge

Theoretical work led us to the view that objectivist assumptions were inappropriate to schemes of intellectual organisation for retrieval. As a more appropriate conceptualisation we began to think of knowledge as socially constructed. The first paper records our first exploration, in March 1972, of the implications of this position for information handling, namely that we should think of relating material to the dimensions by which knowledge comes to be structured. The second paper describes an early attempt to operationalise the notion of dimension. We next began to explore the relation between the notions of dimension and perspective. We were being pressed to give concrete illustration of our thinking at this time, and the third paper represents a premature attempt to crystallise our ideas by organising a sample of material in terms of selected perspectives characterised by means of dimensions. We then returned to more systematic investigation. It should be emphasised that we regard relations between dimensions and perspectives as something to be investigated not imposed. A further paper outlines the steps involved, and the methods appropriate to the questions confronting us. Finally, we have linked together a series of reports and discussion papers associated with in-team meetings which describe our progress to date in this line of investigation, up to the time when we halted it to prepare volume 2 of this report, which uses a portion of our sample to illustrate the point our work has reached at this time. The papers in this set were prepared for internal working use, and we include them here as raw material only. As such, they may serve to illustrate our method. The ideas are now at a stage for discussion with subject experts. We do not feel able as yet to compile a formal report or to give any overview of this phase of work because it is still in evolution. Section IV contains some general comment on the work.

III 6A (i) Suggested principles for reference structure building, May 1971.

1. In the index entries we feel it necessary (in order to convey the problem dealt with in a given item as a whole) to use, German fashion, concept phrases such as perception of role of language. This reduces fragmentation and hence impediment to immediate comprehension (e.g. Language. Role. Perception) because the relationships have to be created as the string is read. However, it seems desirable that the lead terms (i.e. terms selected from those appearing in an entry for use as subject headings) should be basic concepts (e.g. Language in our example) not compound concepts (such as Language Role Perception) which may often be 'created' for the purpose of a given writer but have no permanence. A basic concept may often of course be a multi-word concept.

2. This raises the question of a definition, at least an operational one, of a 'basic concept'. We suggest a distinction, to start with, amongst:

(a) Descriptive terms (i.e. describing the concrete, valid by reason of the observability of differences and similarities amongst objects e.g. grammar school, comprehensive school);

(b) Variables (i.e. quantifiable notions such as age, and also those 'containing' or associated with an accepted basis for operationalisation e.g. race, sex, occupation);

(c) Constructs (i.e. concepts of an analytical nature such as professionalisation which, to be operationalised, can be measured only indirectly and/or for which there are no universally accepted measurement continua).

We feel that these types of concepts need handling rather differently in terms of reference structure.

Some concepts on our list may need to be grouped in further categories e.g. statistical techniques, conceptual frameworks. The present three are considered in further detail.

3. Individual types of concepts

(a) Descriptive terms These are terms such as 'developing countries', 'pre-school children', 'ghetto schools', which may be people, places or concrete things. Additionally terms such as 'secondary education' might be included here as simple grouping concepts for such terms. They may fairly readily be related to each other in hierarchical fashion, though not in mutually exclusive 'trees'.

(b) Variables These include for instance 'age-grade', 'social class', and themselves may be subdivided in a fairly well recognised and accepted pattern. However their relationships to each other are formal rather than relationships in terms of content (except in the temporary context of a given study). Their reference structure will therefore probably most appropriately take the form of a set of discrete mini-taxonomies, quasi-hierarchical and mutually exclusive.

(c) Analytical constructs These concepts are not associated in terms of hierarchies. Rather one such concept may be considered from the aspect of another, e.g. role conflict. As a general rule it may be suggested that when this type of association is commonly denoted by a noun phrase, as in 'role conflict', it may be regarded as a permanent association or compound concept; when common linguistic practice would prefer a prepositional phrase, as in 'perception of role of language', it is probably not to be regarded as an established compound concept nor used as such as an index subject heading. A compound analytical concept, though with apparent concrete associations as in 'teacher role', may not unreasonably be fragmented into the 'basic' elements of 'teacher' (a generalised notion) and 'role'. In this case however it seems important to distinguish between this type of association and that between 'teachers' and 'role' - a mixed (descriptive/analytical) association of concepts. We seem to have in the latter less a permanent conceptual link than an 'occasional' one ('role' also will frequently be used in a non-technical sense). Within the analytical range there are a range of conceptual links in process of becoming established. These may fluctuate in the way they are expressed linguistically (prepositional/noun phrase) and an intelligent guess has perhaps to be made as to their chances of being accepted as 'valid' or permanently useful association.

Perhaps the most difficult problem is that an individual basic concept may be defined in a variety of ways, some highly idiosyncratic. Compound concepts increase the range of individual variations. Ideally, broad patterns of usage should be identified and both basic and associated compound concepts (e.g. role, role conflict, role taking etc.) should be related to their definitions and groups of index entries separated according to conceptual distinctions. Reliable identification of broad types of concept definition, even if only on a simple discipline basis, is not impossible, but classifying the use of a concept by an individual author even in these terms presents enormous difficulties. The knowledge of the other concepts with which a given concept is associated and/or conceptual framework, which the index entry provides, may serve to clarify usage, but this will not invariably be the case.

This essentially remains a problem even when the number of items to which a given label is attached is few.

Coupled with the problem of definitions (essentially a homonym problem) is that of synonyms. There are probably no true synonyms amongst analytical concepts, nor even concepts which may for practical purposes be regarded as such. Thus the nature of the reference structure will probably be composed of overlapping clusters of basic and associated compound concepts, e.g. role and associates overlapping both with position and status and their associates.

4. Relations between types of concepts

Mixed compounds, as distinct from pure compounds (we are having to create our own jargon), may logically be descriptive/variable, descriptive/analytical, variable/analytical or descriptive/variable/analytical. Our list contains for instance 'age-grade consciousness', 'familial socialization', 'school as organisation', 'racial subgroups' etc. In some cases one would wish to refer from the second component to the first as from a general type to a particular case or as bearing some sort of equivalence relationship, but not in every case. We need guiding principles.

Basic concepts of one type might also perhaps usefully be linked to basic concepts of another type e.g. organisation to hospital and other establishments which might be viewed as such. Such concept to concept links however run the risk of referring the user to a certain amount of material in which although a conceptual or operational link might be present, in some cases it would not. These are again temporary associations rather than established ones. As an aide mémoire to the user, where the frequency with which the link is found in the documents warrants it, 'may be' references could be included but the conventional see also reference would not be appropriate.

III 6A (ii) Handling of single cluster of concepts (sample of material broadly related to achievement): report to Working Party

1 Objectives are (in order of priority):

(a) to organise material in a way which does no violence to the ways in which either sociologists or educationalists view work in the field;

(b) to select, from the possibilities which satisfy objective (a), the approach most generally useful and immediately meaningful in relation to the purposes and enquiries for which documents will be sought;

(c) to select, from options meeting objectives (a) and (b), a method which is widely applicable to a range of areas and not merely the one (i.e. achievement) currently being studied.

It is taken as given that a number of 'commonsense' requirements also must be satisfied (e.g. that the principles of organisation should be such that they can be applied consistently, that cross references should guide users from their chosen entry points to the heading(s) under which required material is located, and so on).

2 Stages of work in studying alternative ways of organising items in the achievement sample were:

(a) to list possible kinds of main headings;

(b) to sort the material all ways, noting the problems experienced in each case, and identifying viable options;

(c) to list possible kinds of subdivision;

(d) to apply those, as appropriate (some kinds of main headings preclude some kinds of subdivision), to viable sets of main headings;

(e) to produce sample indexes illustrating different combinations of main and subheadings for the consideration of the Working Party.*

Main and subheadings are discussed separately.

3 Main headings There are three general questions to consider in relation to main headings:

(a) Origin of terms used as headings:
  (i) Author's words
  (ii) Translation of author's words into words with agreed definitions.
  (iii) Translation of author's words into words designated to represent cluster of related notions

---

* The documents used in the exercise have been confined to those in which notions feature which the authors have named 'ability', 'achievement', 'aptitude', 'attainment', 'intelligence', 'performance', 'skill', 'success', 'talent' etc. We recognise however that when, for instance, 'status attainment' is included in the cluster, it does not require a very great intellectual step to the suggestion that, say, 'mobility' should also be regarded as a member of the cluster. Further discussion of 'boundaries' seems desirable.

(b) Alternative types of intellectual organisation of terms used as main headings:

Type	Examples
(On basis of conceptual distinctions amongst terms.)	
(i) Single 'umbrella' heading	'Achievement' + references e.g. Ability see Achievement
or Various combinations	'Ability and Intelligence' 'Achievement and Aptitude' + references e.g. Intelligence, see 'Ability & Intelligence'
or Each term in cluster as heading	'Ability', 'Achievement', 'Aptitude' etc. + cross references from each term to all others in cluster
(On basis of 'content', but ignoring above conceptual distinctions.)	
(ii) Situational distinctions)	'Achievement (i.e. generalised notions)' 'Achievement (i.e. Educational achievement)' 'Achievement (i.e. Occupational achievement)' 'Achievement (i.e. other)' + references e.g. Ability, see 'Achievement (variously defined)'
(On 'formal' basis ignoring both conceptual and situational distinction)	
(iii) Distinction by role	Either 'Achievement (i.e. General treatments)' 'Achievement (i.e. Element in study)' 'Achievement (i.e. Delimiting scope of study)' 'Achievement (i.e. Subjective 'angles')' Or 'Achievement (i.e. Attempts to define concept)' 'Achievement (i.e. General discussion as 'problem')' 'Achievement (i.e. Element in theory)' 'Achievement (i.e. Variable in empirical work)' 'Achievement (i.e. Contextual, e.g. Population studied)'

(iii) Distinction by role (continued) Or (continued)

'Achievement (i.e. Physical setting, e.g. Sample)'
'Achievement (i.e. Subjective notions, e.g. Expectations)'
'Achievement (i.e. Ideology, e.g. Prescriptions)'

(iv) Distinction by level/unit of analysis

'Achievement (i.e. Individual achievement)'
'Achievement (i.e. Group achievement)'

(NB In the cases of (ii) to (iv) some term in the cluster other than achievement might be considered preferable as a label.)

(c) Positioning of related notions in alphbetical sequence:

Kept together, e.g. 'Achievement (i.e. Educational achievement)'
'Achievement (i.e. Occupational achievement)'

Dispersed, e.g. Educational achievement
....
....
Occupational achievement

Experimentation on each of these sets of alternatives (a-c) will be reported separately.

4 Experimentation with authors' v 'standardised' terms

The considerable number of terms in the 'achievement' cluster, and the variant definitions associated with each, mean that, if author's terms formed the subject headings, items would be located in a number of sections of the index. Whilst the material, even if physically separated, could be linked verbally by means of the cross reference structure, the index user would be forced to refer to material under most of the subject headings to locate required material (i.e. matching his definitions rather than his chosen terms). Subject headings must clearly rest upon a more reliable basis.

5 Experimentation with alternative types of intellectual organisation of main headings.

(i) Discussion with the Working Party suggested that it would be difficult to decide upon 'agreed' definitions of individual terms which could be used as the basis for 'standardised' subject headings. The problem is not simply that different definitions of a given term may vary, but that the kind of conceptual distinction made may vary from one writer to another.

For example, viewed in one way 'achievement' and 'performance' may be synonymous in the sense of 'position on a scale'. But it has been pointed out that 'achievement' may additionally be said to contain the notion of 'striving' and is often associated by the sociologist with 'mobility' or by the psychologist with personality factors such as motivation. Some research involves an analytical distintion between 'achievement' as a construct and 'level of achievement' as measured by psychological tests. Taking this view, some researchers take results (performance) on, say, an ability test as an indicator of 'achievement'. Depending on standpoint 'achievement' is thus at one and the same time broader, coextensive and narrower in relation to performance, or perhaps even 'something' quite distinct. To opt for a one term/one definition basis for subject headings would be to misrepresent a large proportion of the literature, and to mislead the index user. It also requires the impossible of the indexer in that there is often insufficient evidence in the document to determine the particular definition an author would have attached to a term had he been making a different kind of distinction (i.e. that which the indexer has in mind rather than the kind he considered appropriate in the context of his work).

It was suggested that by combining closely related notions (e.g. Ability + Intelligence, Aptitude + Achievement) to form subject headings, with cross-references from 'non-used' terms, the number of alternative headings to be searched might be reduced, and reliability increased. All the logically possible combinations for four terms in the above example were tested. The effect was still to separate items which conceptually could not be separated. For example the particular combination suggested above is unacceptable in the light of such studies as one in the sample in which performance on an ability test was taken as an indicator of achievement. Reliability and thus predictability is still lacking in subject headings formulated along those lines.

One is forced to the conclusion that (assuming indexing is to be something more than indexing of words) a document may be indexed as having relevance to a fairly broad area of research interest (a procedure which is of limited practical value to the index user), but that specific concept indexing of the kind described in indexing manuals is not at this time possible in our field.

(ii) Whilst conceptual distinctions of the kind discussed above present apparently insuperable problems for indexing, another kind of distinction suggested by the Working Party has been explored and seems more promising. This is what we have called a 'situational' distinction; it involves 'cutting the cake' (i.e. achievement and related research) in a different way, in terms of the situation actually studied. In the case of our material this is in the main a distinction between research on achievement in the educational and the occupational setting.

Such an approach is not without its problems. A separate heading for 'generalised notions of achievement' can handle studies in which the setting is not particularised, but the distinction between individual/group level study is for instance masked. Nevertheless an educational/occupational split offers far higher promise of reliability (synonyms such as educational/academic performance etc. present little conceptual conflict), and the distinction is likely to be both meaningful and useful to both the educationalist and the sociologist. Those interested in ability as distinct from achievement (in their definition) must needs scan through all the items under the heading for their sphere of concern, but this would be the case even had the split been made according to author's chosen terms or indexer's interpretation of concepts from amongst ability, achievement, aptitude etc.

(iii) A further alternative intellectual basis for subject headings, again regarding the ability, achievement etc. cluster as a single cluster and cutting this 'cake' yet another way, is to focus upon 'formal' aspects of the use of this kind of notion in the document being indexed. This type of distinction was discussed at some length by the Working Party and has been further developed after detailed study of our sample of documents in the course of the current experiments.

This approach, though promising, also has its problems. It seems to group distinctive types of work in a way which will be helpful to many in eliminating material non-relevant to purpose though generally relevant in content, but perhaps favours the researcher at the expense of the non-researcher. The same problem also faces the user interested in a specific concept with scanning whole section of the index but the same comment obtains as in the case of the previous approach.

Perhaps more serious is a risk of variable reliability. Even if the detailed set of eight headings is collapsed to four, the indexer may be forced either to take a study at its face value or to make specialist judgments in indexing according to these distinctions. For example, a sample of negroes may be used in a given study of the relationship of several variables to low achievement. Achievement may be seen as a variable or as delimiting (contextually) the problem under investigation (e.g. if the population is 'low achievers' and negroes are not being studied because they are negroes). Documents do not always provide a clear answer to the indexer attempting to determine the nature of author's problem.

(iv) Distinction by level/unit of analysis is likely to produce broadly a division between sociological analysis and educational/psychological studies of a given topic. It may not be relevant to all the areas of study represented in our index, nor will it necessarily be the first consideration in the index user's mind, even the sociologist user. In the case of certain material (for example that dealing with goals or values), however, it may well be the approach by far the most appropriate to the material.

It is clear that different areas will best be handled in different ways; there is no universally 'best' or 'correct' form of organisation, and complete uniformity is probably less desirable than a limited number of easily recognisable patterns, the pattern being selected to suit work in a given area. There will sometimes be a balance to be struck between the demands of the material and the demands of user approaches. In the case of the present sample the problem does not arise (one item only dealt with achievement at the group level).

6 Physical organisation of main headings in index It seems clear that, whatever the intellectual organisation underlying main headings in the index, there is considerable overlap of interest amongst the documents in the sample, and that, if for this reason alone, a means should be found to locate the sets into which it is grouped physically together in the index rather than to disperse them at various places in the alphabetical sequence. There is also the point that the principle underlying division into sets will not satisfy all users and that those who require subsets of material from several of the main sets will be best served by such a procedure.

7 Subheadings To a large extent the way in which material is sub-divided under main headings can minimize problems of locating material for those whose approach to the area cuts across that adopted for main headings.

Possible principles of subdivision include those suggested as a basis for main headings, as well as others which apply only at the subheading level. The latter may be used either as alternatives to the conceptual, situational, formal or level of analysis principles suggested for main headings or, if two levels of subheadings are required to make large sections of the material manageable, as additional to them. It had been hoped that subdivision could be kept to one level only in order to avoid complexity or appearance of complexity, but our experiments suggest that some further subdivision seems to be unadvoidable.*

The types of organisation considered for subheadings are:

(i) Conceptual

(ii) Situational

(iii) Formal

(iv) Level/unit of analysis

(v) Other major variable(s) or concept(s)

(vi) Population studied (or sample if unclear)

(vii) Location

Further options include use of author's v 'standardised' terms (as for main headings). Sets of subheadings may also be broad (fixed) categories as at the main leading level or specific and enumerative. For example, 'Achievement (i.e. variable in empirical work)' may be subdivided according to other variable(s) stated at the level of independent, dependent and perhaps intervening variable[1], or the specific other variable(s) may be named.

Another consideration is whether, under a given set of main headings (e.g. Achievement defined in terms of formal role), the same principle of subdivision should be adopted for each member of the set or the principle of subdivision should be varied for goodness of fit.

8 Experimentation with alternative types of subheading

(i)-(iv) The problems at the subheading level are very similar to those encountered with main headings.

(v) With regard to 'other major variable(s)' there is the problem of studies including a number of variables (e.g. correlational studies). This means a need for extended use of multiple entry unless a convenient summary term (e.g. family background factors) can be found. Additionally, some studies are not readily analysed in terms of variables. The most serious problem is perhaps that users may well be more accustomed to subdivision of a hierarchical nature in indexes and may take some little time to adjust to a rather different mode of subdivision, even if their enquiries are couched in terms of required combinations of variables.

(vi) 'Population studied' presents the problem discussed earlier of determining the nature of the population to which the study is generalisable where the author does not distinguish clearly between sample and population. Multiple entry is a possible solution, but has the attendant bulk drawback. This principle is of little value in the case of theoretical material.

---

* The italicising of words in document descriptions has been considered as a method of bringing out another factor without altering the pattern of subheadings. This may have a value but this will be mainly for scanning by eye rather than physical organisation of documents, since the descriptions employ authors' words and hence alphabetisation by italicised words would tend to scatter as much as to group related terms.

[1] An objection to this procedure is the danger of appearing to imply causality.

(vii) 'Location' is an obvious principle to consider in subdividing comparative studies. Its more general value is questionable.

In considering the further options noted, the reasoning put forward in discussing authors' v 'standardised' terms holds as much for subheadings as for main headings. The broad v specific/enumerative issue to a large extent depends on distribution of material over categories and the need for categories of manageable proportions. Enumeration, providing the underlying principle governing the procedure is clearly understood by indexers, is not likely to involve any serious loss of reliability. The principle may not however be as immediately obvious to the user as in the case of headings of a more generalised nature, and hence there may be loss of predictability.

9 Conclusions No firm conclusions can be drawn from our experiments, though certain alternatives can probably be ruled out on the reasoning set out in this report. It would be in any case a very great task to reproduce indexes illustrating all the various combinations of headings and subheadings. We have therefore selected those alternatives which seem viable in practical indexing terms (e.g. reasonable degree of reliability) and illustrated them by sample indexes for consideration by the Working Party in the light of the objectives stated at the beginning of this report. (The indexing problems presented by the body of documents thus indexed might well provide material for the paper the group plans to publish.)

The following table summarises our findings in a very general way and indicates the alternatives for which our experimental indexing of the 'achievement sample has been reproduced for circulation. Generalisability to other areas of our field is a matter for discussion

	Main headings			
Subheadings	Conceptual basis	Situational basis	Formal basis	Level/unit of analysis
Conceptual basis	Null category	Ruled out on grounds of definitional conflict		
Situational basis	Ruled out on grounds of definitional conflict	Null category	* (no.3)	priate sample
Formal basis		* (no.1)	Null category	
Level/unit of analysis		Not appropriate to 'achievement' sample		Null category
Other major variable(s) or concept(s)		This principle of main heading illustrated in no.1 + subheadings in no.4.	* (no.4) Using mixture of types of sub-heading	Not appro- to 'achievement'
Population (or sample) studied		* (no.2)		
Location		Not appropriate to 'achievement' sample		

* = sample index produced

It may be of interest to study these experiments further from the point of view of insight into intellectual aspects of the indexing process. By handling different characteristics of division separately, it is possible to compare them both with each other and also with regard to interaction effects.

An interesting aspect of these experiments has been that in indexing there is a need to ask somewhat different questions about a given document depending on the characteristic(s) of division underlying the indexing scheme, though some schemes in the event, in terms of questions to be asked, proved to be more similar than was at first apparent. (eg situation and population). We were conscious of a risk that, having internalised the approach required by a given scheme, the indexer might be influenced to interpret the documents being indexed in the light of the questions the scheme requires him to ask rather than in terms of the questions posed by the documents themselves.

Another point is that, irrespective of the characteristic(s) of division adopted, one might expect the 'sum' of the categories to be the same. This does not seem necessarily to be the case. Depending on the principle used, it was found that 'achievement measurement and testing', for example, seemed sometimes to fit within the available categories and sometimes to fall outside them. This seems to suggest that the 'conceptual space' represented by a given term will vary according to the characteristic of division adopted. Where several characteristics of division are combined in an indexing language (be it a formal classification or alphabetic with cross references) it is possible that the indexer may unconsciously shift from one definition of a term to another.

There seem to be various ways in which the intellectual structure of the indexing language may influence either interpretation or representation of documents, and we have no means of telling whether the effect is to improve or to detract from the quality of indexing. This seems to be an important question to which we should devote further attention.

Sample index to material relating to achievement

(This index covers a range of concepts - intelligence, performance, ability, aptitude etc - amongst which no valid distinctions can be made, and which it is hence unhelpful to separate in an index.)

ACHIEVEMENT (i.e. Educational achievement)

SUBDIVIDED BY ROLE

Element in study

Analysis of relation between achievement of pupil and organizational differentiation in school, distinguishing between vertical/horizontal and elective/selective dimensions of differentiation: effects (direct and indirect) of individual, interpersonal and school variables: bearing on classroom situation. Implications for equality of educational opportunity. 041

Comparison (in relation to survey data) of two models of relation between SES and adolescent achievement: alternative temporal sequences of intervening variables (educational expectations, ambition and intelligence).
Source of data - Reanalysis of data (on male high school freshmen). 033

Comparison of effect of homogeneous/mixed grouping in maintained and direct grant schools on achievement (A-level results) of able pupils, taking account of arts/science specialism; relation between success rates at O-level and A-level.
Source of data - Sample of results of boys from schools (73 maintained and 15 direct grant) in [?England]. 052

Consideration of research into determinants of social class and racial differences in intelligence (and hence achievement), distinguishing between environmental and genetic factors; bearing of argument (that genetic factors are major determinant) on development of educational programmes for disadvantaged children. 060

Correlations amongst family background (size of family, position in family, occupational status of parents, parental interest), intelligence, personality variables (teacher ratings and test scores) and school achivement, taking account of sex.
Source of data - Sample of 4,000 junior school children (age 11) in England. 058

Correlations amongst personality variables, achievement and ability, taking account of sex.
Source of data - Sample of 206 junior school children (age 11) in England. 057

Discussion of Coleman's subcultural model: background to investigation of some relationships amongst adolescent values (as operationalized by Coleman), social participation in school, educational and occupational attainment, with consideration of persistence of adolescent value orientations into adulthood: relation to Coleman's findings.
Source of data - Sample of 343 students from 1 high school graduating in 1962 (follow up 1967) in U.S. Longitudinal study. 068

ACHIEVEMENT (i.e. Educational achievement) cont'd

SUBDIVIDED BY ROLE cont'd

Element in study cont'd

Empirical test and development of model (formulated by Sewell, Haller and Portes) linking social origin and mental ability with occupational and educational status attainment (length of education): relationships amongst intervening variables (educational and occupational aspirations, academic performance and expectations of significant others).
Source of data - Reanalysis of data (on high school seniors). 039

Findings of existing research on association between grade point average and work orientation; also to findings on influence of college experience upon occupational goals of students. Related to comparison of dominant orientation to work in academic and business worlds, using Rosenberg's classification of occupational goals, and taking account of influence of academic discipline.
Source of data - Reanalysis of data (on college bound high school students and college students). 059

Investigation of achievement (educational and occupational) of sons of families receiving welfare compared with that of non-welfare sons from comparable home environment. Test of theory that distinctive opportunity structures confront different economic strata.
Source of data - Reanalysis of data (national sample). 038

Investigation of correlates (sex, field of study, social class background, personality factors, educational values and level of aspirations) of academic success: comparison amongst students who 1. obtained good degree, 2. obtained other degrees, and 3. had not yet graduated.
Source of data - Sample of health and social science students from one university in England. 067

Investigation of indirect influence of SES on academic achievement (ability, school grades and expected amount of future education), taking account of northern/southern differences and sex: personality factors (self-perceptions, perceptions of opportunities and achievement values) as intervening variables.
Source of data - Sample of 2,588 negro high school students in U.S. 010

Investigation of influence of peer group factors and participation in extracurricular activities (as vehicle for interpersonal competition and status acquisition) on student aspirations to attend college, college attainment (total number proceeding beyond first year in college) and goal fulfillment (number of aspiring students proceeding beyond first year), taking account of family SES, student IQ and high school performance.
Source of data - Sample of 297 high school seniors (follow up post high school) in U.S. Longitudinal study. 042

Investigation of influence of sex on attitude towards school, according to type of school attended (taken as a measure of achievement).
Source of data - Sample of 384 students (ages 14-15) from 11 schools (6 grammar and 5 secondary moderns) in [?England]. 062

ACHIEVEMENT (i.e. Educational achievement) cont'd

SUBDIVIDED BY ROLE cont'd

Element in study cont'd

Investigation of interrelationships amongst personality variables, study methods and academic performance (A-level grades and marks at end of first year of college), taking account of sex and type of educational establishment.
Source of data - Sample of 157 students from one university and one college of education in [?England]. 055

Investigation of relation between social origins and career success: college selectivity, college prestige and academic achievement as intervening variables. Contribution to study of continuing (post college career) effects of ascriptive factors on patterns of stratification.
Source of data - Sample of 4,000 engineering graduates in U.S. 029

Investigation of relationship between school attainment and non-intellectual variables (personality, sex and socio-economic background): comparison between high/low ability groups and complete ability range.
Source of data - Sample of 2,538 school children (ages 10-14) in Scotland. 056

Investigation of relative importance of social origins, college grades, amount of formal education, occupational values and size of work organization as factors in career success of graduates. Contribution to study of continuing (post college career) effects of ascriptive factors on patterns of stratification.
Source of data - Sample of 1,360 engineers graduating in 1947-61 in U.S. 028

Investigation of relative importance of traditional ascriptive (social origin) and achievement (academic achievement) factors as compared with 'new' organizational dimensions (college selectivity, college prestige and work organization recruitment) as influences upon career success of graduates. On basis of model of sponsored v contest mobility.
Source of data - Sample of 1,142 engineers and managers. 017

Investigation of relative importance over time of possible determinants (ability, selectivity of college, grade point average, number of years in graduate enrollment, higher degrees, freshman, senior and 1964 prestige expectations) of occupational attainment (1968) of college graduates (7 years after graduating). Implications for understanding relationship between graduation and intergenerational mobility.
Source of data - Reanalysis of data (college graduates in NORC longitudinal survey). 069

Investigation of relative influence of desegregation (at school and classroom levels) and family background (mother's education and possessions in home) on verbal achievement of negro students, with consideration of grouping procedures within schools.
Source of data - Reanalysis of data (on ninth grade students). 064

ACHIEVEMENT (i.e. Educational achievement) cont'd

SUBDIVIDED BY ROLE cont'd

Element in study cont'd

Investigation of relative influence of physical attractiveness and educational attainment on upward mobility through marriage: class variation.
Source of data - Reanalysis of data (on women studied longitudinally in Oakland Growth Study). 009

Notion that level of child's achievement is best predicted from the average age of classmates: basis for investigation of seasonal effect (month of birth), length of schooling and age-group position effect as factors in mathematics achievement: studied in relation to differences in school systems in different countries (Japan, Sweden, England, U.S., France and West Germany).
Source of data - Reanalysis of data (on students, ages 13-14). 051

Survey of educational achievement and post school career of applicants to Birmingham University School of Mathematics with regard to differences by sex and social class.
Source of data - Sample of 1,633 applicants during 1959-62 to one university in England. 014

Survey of housing circumstances of young children; effect on educational performance. Implications for equality of educational opportunity.
Source of data Reanalysis of data (on children, ages 7-8). 066

Survey of research into relationship of socialization and personality factors with performance: purpose to assess empirical and theoretical status of alternative hypotheses (cultural deprivation, culture conflict, and educational deprivation) for understanding academic failure of negro youth; suggestions for further research. 061

Viewed as context

Comparison of effect of homogeneous/mixed grouping in maintained and direct grant schools on achievement (A-level results) of able pupils, taking account of arts/science specialism; relation between success rates at O-level and A-level.
Source of data - Sample of results of boys from schools (73 maintained and 15 direct grant) in [?England]. 052

Investigation of correlates (sex, field of study, social class background, personality factors, educational values and level of aspirations) of academic success: comparison amongst students who 1. obtained good degree, 2. obtained other degrees, and 3. had not yet graduated.
Source of data - Sample of health and social science students from one university in England. 067

ACHIEVEMENT (i.e. Educational achievement) cont'd

SUBDIVIDED BY ROLE cont'd

Viewed as context cont'd

Investigation of relationship between school attainment and non-intellectual variables (personality, sex and socio-economic background): comparison between high/low ability groups and complete ability range.
Source of data - Sample of 2,538 school children (ages 10-14) in Scotland. 056

Subject notions

Thesis that social class of student determines teacher's initial expectations of student academic success: differential effect on teacher-student interaction: growth of mutually accepted classroom stratification system.
Source of data - Sample of one class of black ghetto children (kindergarten through second grade) in U.S. Observational study. 034

SUBDIVIDED BY 'PEOPLE'

Categories of achievers defined by educational/developmental characteristics
Students (general)

Survey of research into relationship of socialization and personality factors with performance: purpose to assess empirical and theoretical status of alternative hypotheses (cultural deprivation, culture conflict, and educational deprivation) for understanding academic failure of negro youth; suggestions for further research. 061

Children (at primary education stage)

Analysis of relation between achievement of pupil and organizational differentiation in school, distinguishing between vertical/horizontal and elective/selective dimensions of differentiation: effects (direct and indirect) of individual, interpersonal and school variables: bearing on classroom situation. Implications for equality of educational opportunity. 041

Correlations amongst family background (size of family, position in family, occupational status of parents, parental interest), intelligence, personality variables (teacher ratings and test scores) and school achievement, taking account of sex.
Source of data- Sample of 4,000 junior school children (age 11) in England. 058

Correlations amongst personality variables, achievement and ability, taking account of sex.
Source of data - Sample of 206 junior school children (age 11) in England. 057

Survey of housing circumstances of young children; effect on educational performance. Implications for equality of educational opportunity.
Source of data - Reanalysis of data (on children, ages 7-8). 066

ACHIEVEMENT (i.e. Educational achievement) cont'd

SUBDIVIDED BY 'PEOPLE' cont'd

Categories of achievers defined by educational/developmental characteristics cont'd

Children (at primary education stage) cont'd

Thesis that social class of student determines teacher's initial expectations of student academic success: differential effect on teacher-student interaction: growth of mutually accepted classroom stratification system.
Source of data - Sample of one class of black ghetto children (kindergarten through second grade) in U.S. Observational study. 034

Young people (at secondary education stage)

Analysis of relation between achievement of pupil and organizational differentiation in school, distinguishing between vertical/horizontal and elective/selective dimensions of differentiations: effects (direct and indirect) of individual, interpersonal and school variables: bearing on classroom situation. Implications for equality of educational opportunity. 041

Comparison (in relation to survey data) of two models of relation between SES and adolescent achievement: alternative temporal sequences of intervening variables (educational expectations, ambition and intelligence).
Source of data - Reanalysis of data (on male high school freshmen). 033

Comparison of effect of homogeneous/mixed grouping in maintained and direct grant schools on achievement (A-level results) of able pupils, taking account of arts/science specialism; relation between success rates at O-level and A-level.
Source of data - Sample of results of boys from schools (73 maintained and 15 direct grant) in [?England]. 052

Discussion of Coleman's subcultural model: background to investigation of some relationships amongst adolescent values (as operationalized by Coleman), social participation in school, educational and occupational attainment, with consideration of persistence of adolescent value orientations into adulthood: relation to Coleman's findings.
Source of data - Sample of 343 students from 1 high school graduating in 1962 (follow up 1967) in U.S. Longitudinal study. 068

Empirical test and development of model (formulated by Sewell, Haller and Portes) linking social origin and mental ability with occupational and educational status attainment (length of education): relationships amongst intervening variables (educational and occupational aspirations, academic performance and expectations of significant others).
Source of data - Reanalysis of data (on high school seniors). 039

ACHIEVEMENT (i.e. Educational achievement) cont'd

Categories of achievers defined by educational/developmental characteristics cont'd

SUBDIVIDED BY 'PEOPLE' cont'd

Young people (at secondary education stage) cont'd

Findings of existing research on association between grade point average and work orientation; also to findings on influence of college experience upon occupational goals of students. Related to comparison of dominant orientation to work in academic and business worlds, using Rosenberg's classification of occupational goals, and taking account of influence of academic discipline.
Source of data - Reanalysis of data (on college bound high school students and college students). 059

Investigation of indirect influence of SES on academic achievement (ability, school grades and expected amount of future education), taking account of northern/southern differences and sex: personality factors (self-perceptions, perceptions of opportunities and achievement values) as intervening variables.
Source of data - Sample of 2,588 negro high school students in U.S. 010

Investigation of influence of peer group factors and participation in extracurricular activities (as vehicle for interpersonal competition and status acquisition) on student aspirations to attend college, college attainment (total number proceeding beyond first year in college) and goal fulfillment (number of aspiring students proceeding beyond first year), taking account of family SES, student IQ and high school performance.
Source of data - Sample of 297 high school seniors (follow up post high school) in U.S. Longitudinal study. 042

Investigation of influence of sex on attitude towards school, according to type of school attended (taken as a measure of achievement).
Source of data - Sample of 384 students (ages 14-15) from 11 schools (6 grammar and 5 secondary moderns) in [?England]. 062

Investigation of interrelationships amongst personality variables, study methods and academic performance (A-level grades and marks at end of first year of college), taking account of sex and type of educational establishment.
Source of data - Sample of 157 students from one university and one college of education in [?England]. 055

Investigation of relationship between school attainment and non-intellectual variables (personality, sex and socio-economic background): comparison between high/low ability groups and complete ability range.
Source of data - Sample of 2,538 school children (ages 10-14) in Scotland. 056

ACHIEVEMENT (i.e. Educational achievement) cont'd

SUBDIVIDED BY 'PEOPLE' cont'd

Categories of achievers defined by educational/developmental characteristics cont'd

Young people (at secondary education stage) cont'd

Investigation of relative influence of desegregation (at school and classroom levels) and family background (mother's education and possessions in home) on verbal achievement of negro students, with consideration of grouping procedures within schools.
Source of data - Reanalysis of data (on ninth grade students). 064

Notion that level of child's achievement is best predicted from the average age of classmates: basis for investigation of seasonal effect (month of birth), length of schooling and age-group position effect as factors in mathematics achievement: studied in relation to differences in school systems in different countries (Japan, Sweden, England, U.S., France and West Germany).
Source of data - Reanalysis of data (on students, ages 13-14). 051

Young people and adults (beyond secondary education stage)

Assessment of validity of Achiever Personality Scale (which takes account of non-intellective variables) as alternative to Scholastic Aptitude Test (which stresses verbal aptitude), high school percentile rank or high school quality (percentage of post high school students) as predictor of academic success of disadvantaged college students. Implications for college selection policies.
Source of data - Sample of 218 freshmen, in 2 groups (85% negro), on Opportunity Award Scheme at one university in 1964-1966 (follow up of first group in 1966) in U.S. Longitudinal study. 065

Findings of existing research on association between grade point average and work orientation; also to findings on influence of college experience upon occupational goals of students. Related to comparison of dominant orientation to work in academic and business worlds, using Rosenberg's classification of occupational goals, and taking account of influence of academic discipline.
Source of data - Reanalysis of data (on college bound high school students and college students). 059

Investigation of correlates (sex, field of study, social class background, personality factors, educational values and level of aspirations) of academic success: comparison amongst students who 1. obtained good degree, 2. obtained other degrees, and 3. had not yet graduated.
Source of data - Sample of health and social science students from one university in England. 067

ACHIEVEMENT (i.e. Educational achievement) cont'd

SUBDIVIDED BY 'PEOPLE' cont'd

Categories of achievers defined by educational/developmental characteristics cont'd

Young people (at secondary education stage) cont'd

Investigation of influence of peer group factors and participation in extracurricular activities (as vehicle for interpersonal competition and status acquisition) on student aspirations to attend college, college attainment (total number proceeding beyond first year in college) and goal fulfillment (number of aspiring students proceeding beyond first year), taking account of family SES, student IQ and high school performance.
Source of data - Sample of 297 high school seniors (follow up post high school) in U.S. Longitudinal study. 042

Investigation of interrelationships amongst personality variables, study methods and academic performance (A-level grades and marks at end of first year of college), taking account of sex and type of educational establishment.
Source of data - Sample of 157 students from one university and one college of education in [?England].

Investigation of relation between social origins and career success: college selectivity, college prestige and academic achievement as intervening variables. Contribution to study of continuing (post college career) effects of ascriptive factors on patterns of stratification.
Source of data - Sample of 4,000 engineering graduates in U.S. 029

Investigation of relative importance of social origins, college grades, amount of formal education, occupational values and size of work organization as factors in career success of graduates. Contribution to study of continuing (post college career) effects of ascriptive factors on patterns of stratification.
Source of data - Sample of 1,360 engineers graduating in 1947-61 in U.S. 028

Investigation of relative importance of traditional ascriptive (social origin) and achievement (academic achievement) factors as compared with 'new' organizational dimensions (college selectivity, college prestige and work organization recruitment) as influences upon career success of graduates. On basis of model of sponsored v contest mobility.
Source of data - Sample of 1,142 engineers and managers. 017

Investigation of relative importance over time of possible determinants (ability, selectivity of college, grade point average, number of years in graduate enrollment, higher degrees, freshmen, senior and 1964 prestige expectations) of occupational attainment (1968) of college graduates (7 years after graduating). Implications for understanding relationship between graduation and intergenerational mobility.
Source of data - Reanalysis of data (college graduates in NORC longitudinal survey). 069

ACHIEVEMENT (i.e. Educational achievement) cont'd

SUBDIVIDED BY 'PEOPLE' cont'd

Categories of achievers defined by educational/developmental characteristics cont'd

Young People (at secondary education stage) cont'd

Survey of educational achievement and post school career of applicants to Birmingham University School of Mathematics with regard to differences by sex and social class.
Source of data - Sample of 1,633 applicants during 1959-62 to one university in England. 014

Categories of achievers defined by occupational characteristics

Workers (general)

Discussion of Coleman's subcultural model: background to investigation of some relationships amongst adolescent values (as operationalized by Coleman), social participation in school, educational and occupational attainment, with consideration of persistence of adolescent value orientations into adulthood: relation to Coleman's findings.
Source of data - Sample of 343 students from 1 high school graduating in 1962 (follow up 1967) in U.S. Longitudinal study. 068

Empirical test and development of model (formulated by Sewell, Haller and Portes) linking social origin and mental ability with occupational and educational status attainment (length of education): relationships amongst intervening variables (educational and occupational aspirations, academic performance and expectations of significant others).
Source of data - Reanalysis of data (on high school seniors). 039

Investigation of relative importance over time of possible determinants (ability, selectivity of college, grade point average, number of years in graduate enrollment, higher degrees, freshman, senior and 1964 prestige expectations) of occupational attainment (1968) of college graduates (7 years after graduating). Implications for understanding relationship between graduation and intergenerational mobility.
Source of data - Reanalysis of data (college graduates in NORC longitudinal survey). 069

Survey of educational achievement and post school career of applicants to Birmingham University School of Mathematics with regard to differences by sex and social class.
Source of data - Sample of 1,633 applicants during 1959-62 to one university in England. 014

Professionals (i.e. RG classes I and II)

Investigation of relation between social origins and career success: college selectivity, college prestige and academic achievement as intervening variables. Contribution to study of continuing (post college career) effects of ascriptive factors on patterns of stratification.
Source of data - Sample of 4,000 engineering graduates in U.S. 029

ACHIEVEMENT (i.e. Educational achievement) cont'd

SUBDIVIDED BY 'PEOPLE' cont'd

Categories of achievers defined by occupational characteristics cont'd

Professionals (i.e. RG classes I and II) cont'd

Investigation of relative importance of social origins, college grades, amount of formal education, occupational values and size of work organization as factors in career success of graduates. Contribution to study of continuing (post college career) effects of ascriptive factors on patterns of stratification.
Source of data - Sample of 1,360 engineers graduating in 1947-61 in U.S. 028

Investigation of relative importance of traditional ascriptive (social origin) and achievement (academic achievement) factors as compared with 'new' organizational dimensions (college selectivity, college prestige and work organization recruitment) as influences upon career success of graduates. On basis of model of sponsored v contest mobility.
Source of data - Sample of 1,142 engineers and managers. 017

Categories of achievers defined by personal/social characteristics

General

Consideration of research into determinants of social class and racial differences in intelligence (and hence achievement), distinguishing between environmental and genetic factors; bearing of argument (that genetic factors are major determinant) on development of educational programmes for disadvantaged children. 060

Ability categories

Comparison of effect of homogeneous/mixed grouping in maintained and direct grant schools on achievement (A-level results) of able pupils, taking account of arts/science specialism; relation between success rates at O-level and A-level.
Source of data - Sample of results of boys from schools (73 maintained and 15 direct grant) in [?England]. 052

Investigation of correlates (sex, field of study, social class background, personality factors, educational values and level of aspirations) of academic success: comparison amongst students who 1. obtained good degree, 2. obtained other degrees, and 3. had not yet graduated.
Source of data - Sample of health and social science students from one university in England. 067

Investigation of influence of sex on attitude towards school, according to type of school attended (taken as a measure of achievement).
Source of data - Sample of 384 students (ages 14-15) from 11 schools (6 grammar and 5 secondary moderns) in [?England]. 062

ACHIEVEMENT (i.e. Educational achievement) cont'd

SUBDIVIDED BY 'PEOPLE' cont'd

Categories of achievers defined by personal/social characteristics cont'd

Ability categories

Investigation of relationship between school attainment and non-intellectual variables (personality, sex and socio-economic background): comparison between high/low ability groups and complete ability range.
Source of data - Sample of 2,538 school children (ages 10-14) in Scotland. 056

Disadvantaged

Assessment of validity of Achiever Personality Sclae (which takes account of non-intellective variables) as alternative to Scholastic Aptitude Test (which stresses verbal aptitude), high school percentile rank or high school quality (percentage of post high school students) as predictor of academic success of disadvantaged college students. Implications for college selection policies.
Source of data - Sample of 218 freshmen, in 2 groups (85% negro), on Opportunity Award Scheme at one university in 1964-1966 (follow up of first group in 1966) in U.S. Longitudinal study. 065

Consideration of research into determinants of social class and racial differences in intelligence (and hence achievement), distinguishing between environmental and genetic factors; bearing of argument (that genetic factors are major determinant) on development of educational programmes for disadvantaged children. 060

Investigation of achievement (educational and occupational) of sons of families receiving welfare compared with that of non-welfare sons from comparable home environment. Test of theory that distinctive opportunity structures confront different economic strata.

Source of data - Reanalysis of data (national sample). 038

Races

Assessment of validity of Achiever Personality Scale (which takes account of non-intellective variables) as alternative to Scholastic Aptitude Test (which stresses verbal aptitude), high school percentile rank or high school quality (percentage of post high school students) as predictor of academic success of disadvantaged college students. Implications for college selection policies.
Source of data - Sample of 218 freshmen, in 2 groups (85% negro), on Opportunity Award Scheme at one university in 1964-1966 (follow up of first group in 1966) in U.S. Longitudinal study. 065

Investigation of indirect influence of SES on academic achievement (ability, school grades and expected amount of future education), taking account of northern/southern differences and sex: personality factors (self-perceptions, perceptions of opportunities and achievement values) as intervening variables.
Source of data - Sample of 2,588 negro high school students in U.S. 010

ACHIEVEMENT (i.e. Educational achievement) cont'd

SUBDIVIDED BY 'PEOPLE' cont'd

Categories of achievers defined by personal/social characteristics cont'd

Races cont'd

Investigation of relative influence of desegregation (at school and classroom levels) and family background (mother's education and possessions in home) on verbal achievement of negro students, with consideration of grouping procedures within schools.
Source of data - Reanalysis of data (on ninth grade students). 064

Survey of research into relationship of socialization and personality factors with performance: purpose to assess empirical and theoretical status of alternative hypotheses (cultural deprivation, culture conflict, and educational deprivation) for understanding academic failure of negro youth; suggestions for further research. 061

Thesis that social class of student determines teacher's initial expectations of student academic success: differential effect on teacher-student interaction: growth of mutually accepted classroom stratification system.
Source of data - Sample of one class of black ghetto children (kindergarten through second grade) in U.S. Observational study. 034

Women

Investigation of relative influence of physical attractiveness and educational attainment on upward mobility through marriage: class variation.
Source of data - Reanalysis of data (on women studied longitudinally in Oakland Growth Study). 009

ACHIEVEMENT (i.e. Occupational achievement)

SUBDIVIDED BY ROLE

Element in study

Discussion of Coleman's subcultural model: background to investigation of some relationships amongst adolescent values (as operationalized by Coleman), social participation in school, educational and occupational attainment, with consideration of persistence of adolescent value orientations into adulthood: relation to Coleman's findings.
Source of data - Sample of 343 students from 1 high school graduating in 1962 (follow up 1967) in U.S. Longitudinal study. 068

Empirical test and development of model (formulated by Sewell, Haller and Portes) linking social origin and mental ability with occupational and educational status attainment (length of education): relationships amongst intervening variables (educational and occupational aspirations, academic performance and expectations of significant others).
Source of data - Reanalysis of data (on high school seniors). 039

ACHIEVEMENT (i.e. Occupational achievement) cont'd

SUBDIVIDED BY ROLE cont'd

Element in study cont'd

Investigation of achievement (educational and occupational) of sons of families receiving welfare compared with that of non-welfare sons from comparable home environment. Test of theory that distinctive opportunity structures confront different economic strata.
Source of data - Reanalysis of data (national sample). 038

Investigation of influence of social class origin on success in academic occupation. As basis for comparison of competing theories (social class origin as constraint on development v exercise of talent).
Source of data - Reanalysis of data (on graduate faculty members and Ph.D.s). 007

Investigation of relation between social origins and career success: college selectivity, college prestige and academic achievement as intervening variables. Contribution to study of continuing (post college career) effects of ascriptive factors on patterns of stratification.
Source of data - Sample of 4,000 engineering graduates in U.S. 029

Investigation of relative importance of social origins, college grades, amount of formal education, occupational values and size of work organization as factors in career success of graduates. Contribution to study of continuing (post college career) effects of ascriptive factors on patterns of stratification.
Source of data - Sample of 1,360 engineers graduating in 1947-61 in U.S. 028

Investigation of relative importance of traditional ascriptive (social origin) and achievement (academic achievement) factors as compared with 'new' organizational dimensions (college selectivity, college prestige and work organization recruitment) as influences upon career success of graduates. On basis of model of sponsored v contest mobility.
Source of data - Sample of 1,142 engineers and managers. 017

Investigation of relative importance over time of possible determinants (ability, selectivity of college, grade point average, number of years in graduate enrollment, higher degrees, freshman, senior and 1964 prestige expectations) of occupational attainment (1968) of college graduates (7 years after graduating). Implications for understanding relationship between graduation and intergenerational mobility.
Source of data - Reanalysis of data (college graduates in NORC longitudinal survey). 069

Survey of educational achievement and post school career of applicants to Birmingham University School of Mathematics with regard to differences by sex and social class.
Source of data - Sample of 1,633 applicants during 1959-62 to one university in England. 014

ACHIEVEMENT (i.e. Occupational achievement) cont'd

SUBDIVIDED BY ROLE cont'd

Element in study cont'd

Thesis that segregation means lower occupational achievement for negroes (through lack of access to information about employment opportunities): tested by relating attendance at integrated school to white friendships (leading to knowledge about jobs) and to employment in higher paid (non-traditional) jobs.
Source of data - Reanalysis of data (on negroes, ages 21-45). 053

SUBDIVIDED BY 'PEOPLE'

Categories of achievers defined by educational/developmental characteristics

Students (general)

Thesis that segregation means lower occupational achievement for negroes (through lack of access to information about employment opportunities): tested by relating attendance at integrated school to white friendships (leading to knowledge about jobs) and to employment in higher paid (non-traditional)jobs.
Source of data - Reanalysis of data (on negroes, ages 21-45). 053

Young people (at secondary stage)

Discussion of Coleman's subcultural model: background to investigation of some relationships amongst adolescent values (as operationalized by Coleman), social participation in school, educational and occupational attainment, with consideration of persistence of adolescent value orientations into adulthood: relation to Coleman's findings.
Source of data - Sample of 343 students from 1 high school graduating in 1962 (follow up 1967) in U.S. Longitudinal study. 068

Empirical test and developmental of model (formulated by Sewell, Haller and Portes) linking social origin and mental ability with occupational and educational status attainment (length of education): relationships amongst intervening variables (educational and occupational aspirations, academic performance and expectations of significant others).
Source of data - Reanalysis of data (on high school seniors). 039

Young people and adults (beyond secondary education stage)

Investigation of relation between social origins and career success: college selectivity, college prestige and academic achievement as intervening variables. Contribution to study of continuing (post college career) effects of ascriptive factors on patterns of stratification.
Source of data - Sample of 4,000 engineering graduates in U.S. 029

ACHIEVEMENT (i.e. Occupational achievement) cont'd

SUBDIVIDED BY 'PEOPLE' cont'd

Categories of achievers defined by educational/developmental characteristics cont'd

Young people and adults (beyond secondary education stage) cont'd

Investigation of relative importance of social origins, college grades, amount of formal education, occupational values and size of work organization as factors in career success of graduates. Contribution to study of continuing (post college career) effects of ascriptive factors on patterns of stratification.
Source of data - Sample of 1,360 engineers graduating in 1947-61 in U.S. 028

Investigation of relative importance of traditional ascriptive (social origin) and achievement (academic achievement) factors as compared with 'new' organizational dimensions (college selectivity, college prestige and work organization recruitment) as influences upon career success of graduates. On basis of model of sponsored v contest mobility.
Source of data - Sample of 1,142 engineers and managers. 017

Investigation of relative importance over time of possible determinants (ability, selectivity of college, grade point average, number of years in graduate enrollment, higher degrees, freshman, senior and 1964 prestige expectations) of occupational attainment (1968) of college graduates (7 years after graduating). Implications for understanding relationship between graduation and intergenerational mobility.
Source of data - Reanalysis of data (college graduates in NORC longitudinal survey). 069

Survey of educational achievement and post school career of applicants to Birmingham University School of Mathematics with regard to differences by sex and social class.
Source of data - Sample of 1,633 applicants during 1959-62 to one university in England. 014

Categories of achievers defined by occupational characteristics

Workers (general)

Discussion of Coleman's subcultural model: background to investigation of some relationships amongst adolescent values (as operationalized by Coleman), social participation in school, educational and occupational attainment, with consideration of persistence of adolescent value orientations into adulthood: relation to Coleman's findings.
Source of data - Sample of 343 students from 1 high school graduating in 1962 (follow up 1967) in U.S. Longitudinal study. 068

Empirical test and development of model (formulated by Sewell, Haller and Portes) linking social origin and mental ability with occupational and educational status attainment (length of education): relationships amongst intervening variables (educational and occupational aspirations, academic performance and expectations of significant others).
Source of data - Reanalysis of data (on high school seniors). 039

ACHIEVEMENT (i.e. Occupational achievement) cont'd

SUBDIVIDED BY 'PEOPLE' cont'd

Categories of achievers defined by occupational characteristics cont'd

Workers (general) cont'd

Investigation of relative importance over time of possible determinants (ability, selectivity of college, grade point average, number of years in graduate enrollment, higher degrees, freshman, senior and 1964 prestige expectations) of occupational attainment (1968) of college graduates (7 years after graduating). Implications for understanding relationship between graduation and intergenerational mobility.
Source of data - Reanalysis of data (college graduates in NORC longitudinal survey). 069

Survey of educational achievement and post school career of applicants to Birmingham University School of Mathematics with regard to differences by sex and social class.
Source of data - Sample of 1,633 applicants during 1959-62 to one university in England. 014

Thesis that segregation means lower occupational achievement for negroes (through lack of access to information about employment opportunities): tested by relating attendance at integrated school to white friendships (leading to knowledge about jobs) and to employment in higher paid (non-traditional) jobs.
Source of data - Reanalysis of data (on negroes, ages 21-45). 053

Professionals (i.e. RG classes I and II)

Investigation of influence of social class origin on success in academic occupation. As basis for comparison of competing theories (social class origin as constraint on development v exercise of talent).
Source of data - Reanalysis of data (on graduate faculty members and Ph.D.s). 007

Investigation of relative importance of social origins, college grades, amount of formal education, occupational values and size of work organization as factors in career success of graduates. Contribution to study of continuing (post college career) effects of ascriptive factors on patterns of stratification.
Source of data - Sample of 4,000 engineering graduating in 1947-61 in U.S. 028

Investigation of relative importance of traditional ascriptive (social origin) and achievement (academic achievement) factors as compared with 'new' organizational dimensions (college selectivity, college prestige and work organization recruitment) as influences upon career success of graduates. On basis of model of sponsored v contest mobility.
Source of data - Sample of 1,142 engineers and managers. 017

ACHIEVEMENT (i.e. Occupational achievement) cont'd

SUBDIVIDED BY 'PEOPLE' cont'd

Categories of achievers defined by personal/social characteristics

Investigation of achievement (educational and occupational) of sons of families receiving welfare compared with that of non-welfare sons from comparable home environment. Test of theory that distinctive opportunity structures confront different economic strata.
Source of data - Reanalysis of data (national sample). 038

Races

Thesis that segregation means lower occupational achievement for negroes (through lack of access to information about employment opportunities): tested by relating attendance at integrated school to white friendships (leading to knowledge about jobs) and to employment in higher paid (non-traditional) jobs.
Source of data - Reanalysis of data (on negroes, ages 21-45). 053

ACHIEVEMENT MEASUREMENT AND TESTING

Prediction

Assessment of validity of Achiever Personality Scale (which takes account of non-intellective variables) as alternative to Scholastic Aptitude Test (which stresses verbal aptitude), high school percentile rank or high school quality (percentage of post high school students) as predictor of academic success of disadvantaged college students Implications for college selection policies.
Source of data - Sample of 218 freshmen, in 2 groups (85% negro), on Opportunity Award Scheme at one university in 1964-1966 (follow up of first group in 1966) in U.S. Longitudinal study. 065

The following two items would be indexed under the heading 'Achievement motivation' rather than 'Achievement'.

Data from investigations of individual achievement motivation and collective action of disadvantaged negro youth. Used in analysis of concept of internal/external control, distinguishing between self/other responsibility and individual/system blame dimensions for purpose of explaining motivational dynamics of disadvantaged (by minority and/or economic status).
Source of data - Reanalysis of data (on negro students, high school dropouts and adults). 012

Exploration of initial effects of a school desegregation program on negro childrens' achievement motivation (behavioural measures of autonomous and social comparison achievement motivation); influence of school settings with different racial balances. Implications for study of effect of change in reference group norms on personal motivation.
Source of data - Sample of 993 elementary school children(kindergarten - 5th grade) studied in 1965 before desegregation (follow up 1966 after desegregation) in U.S. Longitudinal study. 070

III 6A (iii) [Overall structure of index]

Proposed principles for indexing: background paper

1. Subject headings as frameworks

Those to whom the term 'framework' signifies 'conceptual framework' may consider it inappropriate to apply the term to sets of subject headings. This is in some ways an arid debate, but it seems important to consider some desirable characteristics of a set of subject headings as the means by which documents are pigeon-holed for future reference as required. The individual may then decide for himself whether, in relation to the information storage/retrieval process, subject headings have anything in common with a conceptual framework.

(a) Structure. Irrespective of presentation factors (alphabetical or classified sequence of headings), a set of subject headings constitutes a framework by being structured; explicitly or implicitly, terms (and the groups of material they serve to 'label') are treated as having varying degrees of relatedness, ranging from virtual synonymy to virtual discreteness. Relatedness is conveyed as much by verbal linkage (cross-references) in an alphabetical arrangement as by physical juxtaposition in a classified presentation. By either procedure material is categorised on the basis of relatedness, the sum of the content of the various categories representing the total 'universe' being indexed. One has thus a 'systems' effect, in that every category is a 'sub-system' of some other category or categories.

In many indexing and classification schemes the relationships amongst subsystems within systems, and amongst systems within the universe, are considered to be of a hierarchical (genus/species) nature. In others terms are grouped into mutually exclusive categories called 'facets' (eg 'educands'). Whether relationships are hierarchical, or coordinate, this seems to concern the same dimension of structure, in that the majority of schemes seem to assume that the appropriate structure may be created by representing relationships amongst the 'real' referents of

Cont'd....

these terms. Analytical notions are consequently awkward to accommodate within the pattern. Equally, however, in schemes in which analytical notions and their interrelationships are taken as the basic structure, empirical referents as such have no 'natural' place within the scheme. This problem seems to be associated with a different dimension of the organisation of subject headings from the broader/coordinate/narrower basis of categorisation.

It is widely recognised that by opting for categories formed by applying one characteristic of division, material related in terms of some alternative characteristic will be dispersed (ie distributed categories). Computer-stored indexes can handle complex structures (multiple lattice or tree structures) but in the case of printed page indexes both a concern to avoid undue complexity and considerations of economy usually require that a single set of categories be adopted, additional sets of categories being merely sketched in by chained or branching cross-references. In this way the user is enabled to locate and himself gather together subsets of material forming any given distributed category, even though the convenience of having the material immediately accessible at one place in the index is lacking.

The effectiveness of this procedure depends upon subsets summing to the distributed category. It would appear to be of doubtful value as a means of overcoming the problem of meshing together, within a single system or structure of subject headings, both terms with direct empirical referents, and terms representing analytical notions. The effect may be expected to go beyond inconvenience to some users in locating material because 'their' category has been dispersed. Subsets of categories within a universe structured in terms of 'real' referents will not recombine and sum to categories which are analytical of (ie abstractions from rather than in a whole/part relationship) the 'empirical world', nor does the reverse process seem to hold.

In the case of our index it seems that the content of a subset ('coeducation', say) may be expected to differ depending upon whether it is constructed to form a subset of one broader heading or another (say 'education policy' or 'differentiation process'). It is true that coeducation as a phenomenon may be regarded by definition as differentiation (in some sense an equivalence relationship), but may also

Cont'd....

be otherwise defined, and as studied in a given research may offer no understanding of differentiation processes. Equally studies of differentiation will not necessarily be relevant to all the aspects from which users may study coeducation, particularly non-sociologist users. Thus subsets of material dealing with coeducation, streaming etc will not sum to 'differentiation' if such a category is distributed in favour of primary headings with direct empirical referents. Nor will a combination of subsets of material appearing under various 'analytical' headings sum to a heading such as 'educational policy', were this a distributed category.

Either way a proportion of users will not be able to locate required sets of material at all. This suggests that we have to deal with alternative and overlapping or possibly discrete 'systems' rather than a single system in which subsystems are variously interrelated. In this sense also, therefore, sets of subject headings may be termed 'frameworks'.

To a certain extent an indexer may work initially by 'feel' in considering what documents the user would expect to find grouped together and under what heading, and what groups of material may additionally be relevant. He may later work on the basis of precedent (ie the way in which documents featuring under a given heading contextually define that and related headings). But groups of users are likely to differ in their expectations. To allow for different approaches, and to ensure consistency in indexing, there is no real substitute for clarifying and making explicit general principles upon which relatedness is, for the purpose of compilation and use of our index, to be defined.

(b) Perspective. Our problem seems to be that whilst in one sense all users are alike concerned with the same 'real' universe, universes as studied differ, for instance from one discipline to another. Thus even if boundaries of universes coincide (is this the case with education and sociology?), different preoccupations mean that we are concerned with alternative ways of 'cutting the cake', in other words ways of defining (or imposing structure upon) the same universe. Even where different disciplines apply the same word to the same aspect of the real world (eg authority) it seems unwise to take for granted that the object of study is the same. It is assumed that it is the diversity of objects of study with which we

Cont'd....

are concerned, not simply with words nor with referents to these words which are in any absolute sense 'real'. We are intending to use broad headings so that relationships amongst 'things' resulting from 'cutting the cake' in any single way concern us little. The distributed category principle in any case seems to handle this situation. Our need is to characterise the alternative universes of study we are concerned with, as a basis for determining the relationship between 'things' defined in the process of 'cutting the cake' in one way and 'things' defined by cutting it in any other. Appropriate labels have also to be agreed.

Any simple characterisation of difference in thinking about the 'real world of education' will clearly be a gross oversimplification. Within sociology, 'hardline' sociology may be perhaps equated with the group and institutional levels of analysis; 'soft line' sociology may be held to include analysis in which individuals are considered in terms of a given characteristic such as social class, even if not studied in relation to their functioning as a group. Apart from level of analysis, analysis may be related to different organising concepts (eg conflict). Again some would regard a structure/process distinction as fundamental, whilst others would regard structure as 'process studied at a point in time'. Within education, the buildings, people (and associated practices) may be variously regarded in terms of age level, source of finance, subject of the curriculum etc., and to a certain extent these also cross-cut each other. A broad material/analytical distinction appears to be helpful as a starting point.

2. Link between structure of headings and structure of descriptions.

Our techniques for structuring individual descriptions were designed in the first instance, on the basis of our experience with PRECIS, to meet the general principle that a given heading should be followed by its full context, and that subsequent parts of the entry should be presented in correct relationship to what had gone before.

Individual descriptions, again by virtue of what we had learnt from PRECIS, were not fragmented into separate 'concepts' but divided into much larger 'idea units' to avoid 'garbling'. In manipulation of descriptions to produce index entries meaningful relative to entry point (ie the full context principle), these units were handled as units, the links between these units being the only permitted articulation points.

Cont'd....

The model upon which descriptions were based provided, apart from certain fixed format elements (eg source of data), for two 'moveable' elements – statement of theoretical content, and statement of empirical content, with a mechanism (borrowed from PRECIS) for differential wording of link between those two units depending upon desired order of presentation. (Greater elaboration, providing it was within this basic pattern, was considered possible if it later proved desirable.) It was assumed that, as similarly in PRECIS, terms from both kinds of unit would be required as entry points, though standard terms (rather than author's words as in descriptions) would be used as headings. When a term from an empirical unit featured as a heading, that unit would feature first in the description as context, and a theoretical unit would feature first when a term from such a unit featured as a subject heading. (Any description might of course contain only one or other type of unit rather than both; the description in such cases should however convey the nature of the work.)

These hangovers from PRECIS, coupled with the idea that our material could not be appropriately indexed within a single framework of subject headings, led to the association of the structure of descriptions with alternative frameworks of subject headings, both seen to be associated with possible user approaches. Since users are often both producers and consumers of academic papers it might be expected that an approach considered to provide an appropriate characterisation of documents would also underlie users' formulation of their search problems. To take the example used in the previous section, a given description might contain terms referring, in different elements of the description, to two ways of characterising the same 'real' phenomenon eg streaming/differentiation. Such a description would be indexed under 'differentiation' not because streaming is held to be by definition a form of differentiation but because the author explicitly indicates that he views it as such. Another description, where the author does not so view streaming, would be indexed only under 'streaming'. The empirical/theoretical basis for defining elements of a description, was thus crudely equated with a proposed 'real world'/theory basis of alternative frameworks of subject headings.

Cont'd....

There is no technical link involved but a major point here is that, irrespective of other considerations, if all our headings are pitched at a high level of generality (eg streaming subsumed under differentiation), there is no need for manipulation of descriptions relative to entry point. All terms in descriptions required as entry points would be 'raised' to a subject heading such that the theoretical unit of a description would automatically be required as the first element in an index entry. (The validity of 'raising' a term such as 'coeducation' to a sociological concept such as 'differentiation' seems in any case questionable.)

It is possible that differential indexing according to framework may be considered appropriate, but it seems that the basis upon which frameworks were provisionally defined may be changed for something less simplistic or more appropriate. In this case, if differential structuring of entries relative to entry point is an important consideration, and if the principle of following the entry point by its context is to be retained, then the different ways in which terms are defined (in relation to the kind of framework to which a notion is said to 'belong') will guide redefinition of what is required as appropriate context, and hence the kind of units from which the description will be built up (as well as the kind of links which will be required to avoid garbling when the description is restructured according to entry point).

Attention is now being focussed on the principles governing establishment of subject headings. Depending on decisions on this issue, the possibilities of linking our descriptions to them via our manipulation techniques (subject to non-distortion of documents) will be explored again. But as a matter of policy the availability of our entry manipulation techniques will not be allowed to influence decisions on subject heading organisation.

VW/RAC
5.1.72.

Draft schedule of categories of terms used in indexing achievement and social class samples (some sections rounded out with additional terms)

I 'REAL WORLD' (ie Level 1 on worksheets)

(i) GEOGRAPHICAL LOCATIONS

By country (to be stated unless GB or US)

(ii) FORMAL EDUCATION FRAMEWORK

(a) ESTABLISHMENTS

Education systems (social and administrative aspects)

Schools (admin. categories) -

- Nursery schools and playgroups
- Primary and elementary schools
- Secondary and middle schools

Universities and colleges (admin. categories)

Special kinds of establishment (eg Special schools)

(b) RESOURCES, PROCEDURES, etc, eg:

- Curriculum
- Discipline
- Extracurricular activities
- Grouping
- Mass media
- Selection (etc etc)

? Subgroup when more terms collected.

(c) PROPER NAMES

Tests, schemes etc included selectively

(iii) SOCIAL BACKGROUND TO EDUCATION

Political, economic etc features (not yet considered in detail)

(iv) PEOPLE (individuals or aggregations)

(a) DEFINED BY COMBINED EDUCATIONAL/SOCIAL CHARACTERISTICS

General. Students (youth or adults)

1. Infants and children (at pre-education stage)
2. Children (at primary education stage)
3. Young people (at secondary education stage)
4. Young people and adults (beyond secondary education stage)

(b) DEFINED BY SPECIFIC EDUCATIONAL CHARACTERISTIC, eg:

By Ability

Educational personnel
- Academic staff
- Administrative staff

Dropouts

(c) DEFINED BY SPECIFIC SOCIAL CHARACTERISTIC, eg:

- Disadvantaged
- Parents
- Races
- Women

(v) SOCIAL BACKGROUND TO PEOPLE, eg:

- Environment
  - Home
  - Situational variables
  - Teaching-learning situation
  - Work situation

(vi) SOCIAL/PSYCHOLOGICAL CHARACTERISTICS

(a) DEMOGRAPHIC VARIABLES*, eg:

- Age
- Birth order
- Career
- Marriage
- Physical appearance
- Politics
- Race
- Sex
- Social circumstances, eg:
  - Disadvantagement
- Social class
  - Education (length or level reached)
  - Income
  - Occupation

* Any of these variables may be indexed as of interest per se, but if taken as an indicator, the concept thus operationalised should be indexed rather than the indicator itself.

(b) INDIVIDUAL/GROUP VARIABLES, eg:

- Ability
- Achievement
- Achievement motivation
- Aptitude
- Development phases
  - General
    1. Infancy (0-2 years)
    2. Childhood (3-10 years)
    3. Adolescence (11-17 years)
    4. Adulthood (18+ years)
- Failure
- Genetic characteristics

- Language
- Maturation
- Motivation
- Personality
- Self concept
- Significant others

(c) EDUCATIONAL VARIABLES, eg:

- Study habits
- Subject specialism
- Teaching style

The following will also be used as 'reflexive' subheadings under the 'object' to which they are directed:

- Aspirations
- Attitudes
- Beliefs
- Choice
- Concepts held
- Expectations
- Perceptions
- Predispositions
- Preferences

(An alternative basis for categorization in section I(vi) would be achieved/ascribed characteristics, but this is felt less likely to afford helpful grouping of documents.)

II SOCIOLOGICAL CONCEPTS AND THEORY (ie Level 0 on worksheets)

(i) SOCIAL ORGANIZATION

(a) SOCIAL UNITS (groups or systems) - primary and secondary

- Family
- Small group
- Group by specific characteristic, eg:
  - Minority group
  - Occupational group
  - Peer group
  - Reference group
  - Social class group
- Neighbourhood
- Community
- Organization
- Population
- Society

(b) EDUCATIONAL ESTABLISHMENTS AS SOCIAL UNITS

School (as social unit) - primary and secondary
College (as social unit) - including university

(ii) SOCIAL PROCESSES AND SOCIAL BEHAVIOUR

(a) CHARACTERIZED BY KIND OF ACTOR OR CONTEXT

Social relations (general)

  Familial relations
  Racial relations
  Work relations

(Social relations within educational context)

  Classroom relations
  Home-school relations
  Student interrelations

(b) CHARACTERIZED BY NATURE OF PROCESS

Alienation and anomie

Bureaucracy

Cohesion

Collective action

Communication

Conflict

Control

Democracy

Friendship patterns

Laissez-faire

Mobility

Social change and innovation

Social integration

Socialization

Stratification
  Forms of stratification, eg:
  Opportunity structure

Teaching-learning process
  Preschool education
  Primary and elementary education
  Transition to secondary education
  Secondary and middle school education
  Higher and further education
  Adult education and retraining
  Special forms, eg Compensatory education
(NB All viewed as a process)

(iii) ANALYTICAL CONSTRUCTS

Authority
Culture
Freedom
Goals
Leadership
Norms
Power
Prestige and quality
Role
Selectivity (ie exclusiveness)
Status and position
Value orientations

(iv) THEORETICAL FRAMEWORKS

(a) DIMENSIONS, CONTINUA, PARADIGMS etc, eg:
Authoritarian - democratic
Local - cosmopolitan

(b) TYPES OF THEORY, eg:
Conflict theory
Consensus theory
Environment v Heredity

(v) METHODOLOGY

(a) APPROACHES, eg:
Cross-cultural
Longitudinal
Sociometric

(b) PROBLEMS, eg:
Response bias

Notes on subheadings under terms in Categories I and II

A permanent sub-heading, applicable to any term in Category II, will be 'Definition as concept'. (This should not be confused with the 'reflexive' subheading which may also be used in connection with any of these terms (cf Social class as concept) with Social class + 'reflexive' subheading (ie individuals' concepts of class).)

Both these subheadings will be additional to differential systems of subheadings for the subsidiary sections in each category. The following suggestions as to principles of subdivision are for discussion:

Type of term	Suggested principles for subheadings (following a 'general' section) under each term according to type
I REAL WORLD	
(i) GEOGRAPHICAL LOCATION (Country)	By institution (Education, Economy etc.)
(ii) FORMAL EDUCATION FRAMEWORK	
(a) Education systems (social and admin. aspects)	By authority concerned and/or type of financial support, eg Public/Private; Direct grant/Maintained etc.
Schools (admin. categories)	
Nursery schools and playgroups Primary & elementary Secondary & middle	By type, eg Comprehensive schools
Universities & colleges	By type, eg Colleges of Education
(b) RESOURCES, PROCEDURES etc	Unclear as yet
(c) PROPER NAMES	No subdivision required
(iii) SOCIAL BACKGROUND TO ED.	By institution - political, economic etc.
(iv) PEOPLE	
(a) Defined by ed/soc characteristics (b) Defined by specific ed. characteristics (c) Defined by specific social characteristics	By social-psych characteristic
(v) SOCIAL BACKGROUND TO PEOPLE	By way in which environments are viewed: physical amenities, social situation etc.
(vi) SOCIAL-PSYCH CHARACTERISTICS	
(a) Demographic variables	By people or group
(b) Individual/group variables	By people or group
(c) Ed. variables	By people or group
II SOCIOLOGICAL CONCEPTS AND THEORY	
(i) SOCIAL ORGANISATION	
(a) Social units	By major concepts used in research analysing each one
(b) Ed. establishments as social units	By major concepts used in studying ed. establishments in this way, eg as an organisation.
(ii) SOCIAL PROCESSES AND SOCIAL BEHAVIOUR	
(a) Characterised by actor or context	Unclear as yet
(b) Characterised by nature of process	By social unit (except Teaching-learning by level of education)
(iii) ANALYTICAL CONSTRUCTS	By social unit
(iv) & (v) THEORETICAL FRAMEWORK & METHODOLOGY	Unclear as yet

III EXAMPLES OF SOME ADDITIONAL CLUSTERS CUTTING ACROSS THOSE OF CATEGORIES I AND II

(i) INTRA-LEVEL CLUSTERS

(a) WITHIN 'REAL WORLD' (Category I), eg:

I 'People'; Soc-psych variable
Dropouts; Failure

I 'People'; Soc-psych variable; Ed. procedure
Ability categories; Ability; Grouping

(b) WITHIN SOCIOLOGICAL THEORY (Category II), eg:

II Social unit; Social process
Organization; Bureaucracy

II Social unit; Social process; Analytical construct
Social class group; (Mobility; Status and position
(Stratification

(ii) INTER-LEVEL CLUSTERS

'REAL WORLD' (I) - SOCIOLOGICAL THEORY (II), eg:

I Ed. procedure; II Social process
Grouping; Stratification

I 'People'; Soc-psych variable; II Social unit
Races; Race; Minority group

I 'People'; Soc-psych variable; II Social process
Disadvantaged; Disadvantagement; Opportunity structure

I Soc-psych variable; II Social unit
Age; Peer group

I Soc-psych variable; II Social process
Language; Communication

I Soc-psych variable; Social situation; II Social unit; Social process
(Occupation; Work situation; Occupational group; Work relations
(Career

I 'People'; Soc-psych variable; Social situation; II Social unit; Social process
(Parents; (Marriage; Home; Family; Family relations
(Women (Birth order

I 'People'; Soc-psych variable; II Analytical construct; Theoretical dimension
Academic staff; Teaching style; Leadership; Authoritarian-democratic

Notes:

A given term may, of course, appear in more than one cluster.

Such clusters are formed on the basis of the content of terms rather than, as in Categories I and II, on the basis of 'formal' relationships.

/The.....

The existence of inter-level clusters may be indicated:
EITHER with entries physically separated but verbally linked by cross-references, eg:

Grouping (ie educational procedure)
  Entries under 'grouping'
  See also Stratification
Stratification
  Entries under 'stratification'
  See also Grouping

OR by locating the entries physically together under the more 'general' term, and referring from the 'non-used' term, eg:

Stratification
  Entries for general studies of stratification
 Grouping
  Relevant entries
 Other subheadings
  Relevant entries

Plus reference:
Grouping (or other subheading)
  See Stratification

In either case the user will be enabled to locate required material, but possibly by an indirect rather than a direct route. Expected entry point in relation to expected scope of search (will those seeking material on stratification also be interested in material on grouping and vice versa?) seems to be an important question to consider. Differential handling may be appropriate according to the degree of overlap amongst terms. For example, in the case of the 'Family, Family relations, Home etc.' cluster, these terms and others might well form subheadings under a broad heading such as 'Family and kinship', with references from 'non-used' terms either to the main heading, or to main + subheading as in the International Encyclopaedia of the Social Sciences.

A similar principle might be applied to intra-level clusters where there is definitional overlap of a somewhat different kind. A possible guiding principle is that the entries physically grouped together should relate to documents likely to be used together. Cross-references to further material would then guide the user to items of potential but not assured interest relative to the type of entry point he has selected (ie defined 'formally' as in Categories I and II).

The most relevant practical consideration is perhaps that of overall index bulk. The adoption of broad headings such as 'Family and kinship' or 'Role, status and position' would undoubtedly reduce the amount of multiple entry otherwise required to take account of the problem of definitional overlap*.

/(A.....

---

*We have agreed that it is not for us to attempt to resolve conceptual conflicts. Our task is rather, with some understanding of the nature of the conflicts that exist, to devise practical procedures to enable users to locate relevant documents despite conceptual conflicts (author-author, user-user, author-user).

(A 'contents list' of such headings, as well as detailed indexing in alphabetical sequence of 'non-used' terms, might be provided to aid users.) On the other hand cross-level clusters seem unlikely to be invariably helpful (at the present time at least), so that to this extent at least we need perhaps to differentiate between, say, a 'Parents, Marriage and Home etc' cluster on the one hand and a 'Family, Familial relations' cluster on the other. (A given document would of course probably deal also with other variables or concepts and might be assigned two or three entry points at each level, or rather a subheading under two or three such entry points.) The items in the two samples we are studying may enable us to assess the relative merits of alternative procedures.

III 6A (iv) Establishment of subject headings: further considerations of principle

1. In accordance with the principle, argued in a background paper (see paper III 6A (iii), , that it is inappropriate to organise our material within a single framework of subject headings, we have attempted to characterise and to test alternative combinations of frameworks in relation to our material.

2. The objective has been to devise a set of procedures which will ensure that all the major elements of the problem or problems with which a given document deals will be represented somewhere in the index.

3. Further, descriptions of all the documents dealing with a particular 'thing' should ideally be directly even if not immediately accessible. By this is meant that those interested in a given X may access documents relating to X either by means of a main heading X, by being referred on entry at X to an alternative label for the same notion (X, see Y = X), or being referred to a subdivision of some more general heading or headings (X, see subdivision X under main heading Y...). The latter would be a last resort since we have preempted subdivision for another purpose. The requirement is that by following the reference at any chosen entry point the user should be enabled to locate X, all of X and no more than X. One should add the proviso that X may be a simple concept or it may have to be a conceptual cluster if, as treated in the literature, individual concepts lack conceptual distinctness (cf Achievement etc.).

4. This takes one straight to the heart of our problem. What 'things' are our users interested in? What 'things' do academics write about?

5. At one time it seemed that a major principle might be to decide to work on the basis of either the way in which a document would be viewed by users or the way in which the writer viewed his work. It is now clear that this would leave the basic problem untouched, that is to say the problem of characterising the alternative viewpoints which either users or writers may hold and to which the indexing should be sensitive.

6. There are several general principles which, whatever the approach, seem fundamental:

   (a) a document may deal with more than one problem;

   (b) consistency in breaking a problem down into its elements is required;

   (c) any given element in a problem, however varied the considerations taken into account in making the decision, should be represented at one point in the index only;

Cont'd....

6 cont'd

(d) it should moreover be represented consistently, as further problems 'containing' this same element occur, at the same point in the index;

(e) working definitions of 'sameness' concern more than use of the same word or of effectively synonymous terms; 'sameness' involves consideration of word plus its context.

7. Specifically three alternative though related distinctions seemed to emerge from Working Party discussion:

(a) Real world / theory.

(b) Setting-specific / non-setting-specific.

(c) Non-sociological / sociological.

To these may be added another:

(d) Observables (or directly measurables) / Constructs.

We have experimented with various schemes based on such definitions. Our conclusion is that indexing guided by the logic of the sets of questions generated by any of these distinctions is inadequate in that not all the significant elements of a problem (however minimal the notion of 'problem') can be made directly accessible at the main heading level. (We have agreed that subheading principles are to be related to dimensions of search other than content of terms.)

8. This would suggest one of the following:

(a) we are asking the wrong kind of questions (it has been noted that all these distinctions, whether worked out in terms of broadly methodological considerations, of questions regarding the nature of the referents of the terms used by authors, or of discipline considerations, are closely related);

(b) we are oversimplifying (for instance, we may need to take account of more than one set of questions in identifying the major elements of a problem).

9. The Working Party tended to support the kind of distinction which was proposed, however (operationally) it was arrived at, namely a distinction between concrete or material 'things' and 'things' identified as such by a process of analysis or abstraction from the material world (the realist/nominalist argument is by-passed by thinking in terms of different kinds of objects of study).

10. It would seem therefore that our lack of success should be attributed to oversimplification. Whilst avoiding unnecessary complexity, we must clearly accept as much complexity as is required for reasonably effective treatment of our material.

Cont'd....

11. There are two possible areas in which oversimplification may have occurred:

(a) in specifying the frameworks with which we have to deal (to do justice to either user or author viewpoints);

(b) in specifying valid ways of handling documents in relation to the frameworks as defined.

12. Given that a broad material/analytical basis is appropriate (we have been able to discover no other with which to experiment), it has been assumed that frameworks defined in this way would be regarded as discrete. Whilst both represent (more or less) the same universe of study, they define and present the same total universe in a different way and hence offer alternative and cross-cutting total representations of the same universe and not alternative but partial and complementary subsections of the same universe. Thus one may argue that a problem in a given document should be classified as belonging to one framework or another, and indexed accordingly.

13. We assume provisionally that the frameworks as defined appropriately characterise the alternative perspectives guiding the ways in which users may define their research problems, or view documents in relation to these problems. But these frameworks, even whilst analytically discrete, may be used in a way which is not mutually exclusive. Indexing procedures also should perhaps allow of similar movement from one framework to another.

14. At the same time, use of one, other or both frameworks cannot be left to the discretion of the individual indexer because of the risk of cross-classification (one indexer might decide that a given element of a subject could validly be associated with one framework, another for different reasons might decide that the same element was most appropriately subsumed under a heading in another). The term + context principle cannot be abandoned. Guidelines should therefore specify the circumstances determining with which framework a given element of a subject may legitimately be associated.

15. In this way the discreteness of the frameworks may be preserved; this seems essential to bring together like (and only like) objects of study with like. At the same time, by recognising 'context' as a more complex notion than hitherto, it may be possible to achieve the greater flexibility we seem to need.

Cont'd....

16. The other major alternative line for exploration would seem to require that frameworks be regarded as overlapping rather than mutually exclusive, but that documents be inflexibly classified into one framework only. The logic of our argument with regard to subject headings as frameworks does not suggest that such a basis would be appropriate. It also opens the door to greater risk of inconsistency in indexing in that any guidelines for associating a document with a particular framework are likely to require the exercise of much more subjective judgments.

17. For this reason this latter alternative has not been explored in detail. This does not mean that ideally a variety of analytical frameworks could not be applied to sociological material. But such frameworks would still be regarded as alternative in the sense of cross-cutting rather than alternative in the sense of complementary or overlapping. Also in practice the complexity might be too great for the operational indexing situation.

18. The reasoning in this paper, leading from that in background paper paper III 6A (iii), suggested that the objective should be to develop procedures allowing of the following steps in handling any given subject (a document may deal with more than one subject or problem):

(a) decide whether main focus is empirical or theoretical;

(b) break down subject into (i) central problem, (ii) factors by which study of problem is delimited by authors;

(c) central problem to be classified into categories within framework associated with main focus, each variable as specified by author 'rounded up' to agreed broad heading within appropriate category (special rules needed for handling 'compound' concepts such as 'reflexives');

(d) 'delimiting' factors handled similarly but in less detail according to whichever focus is appropriate;

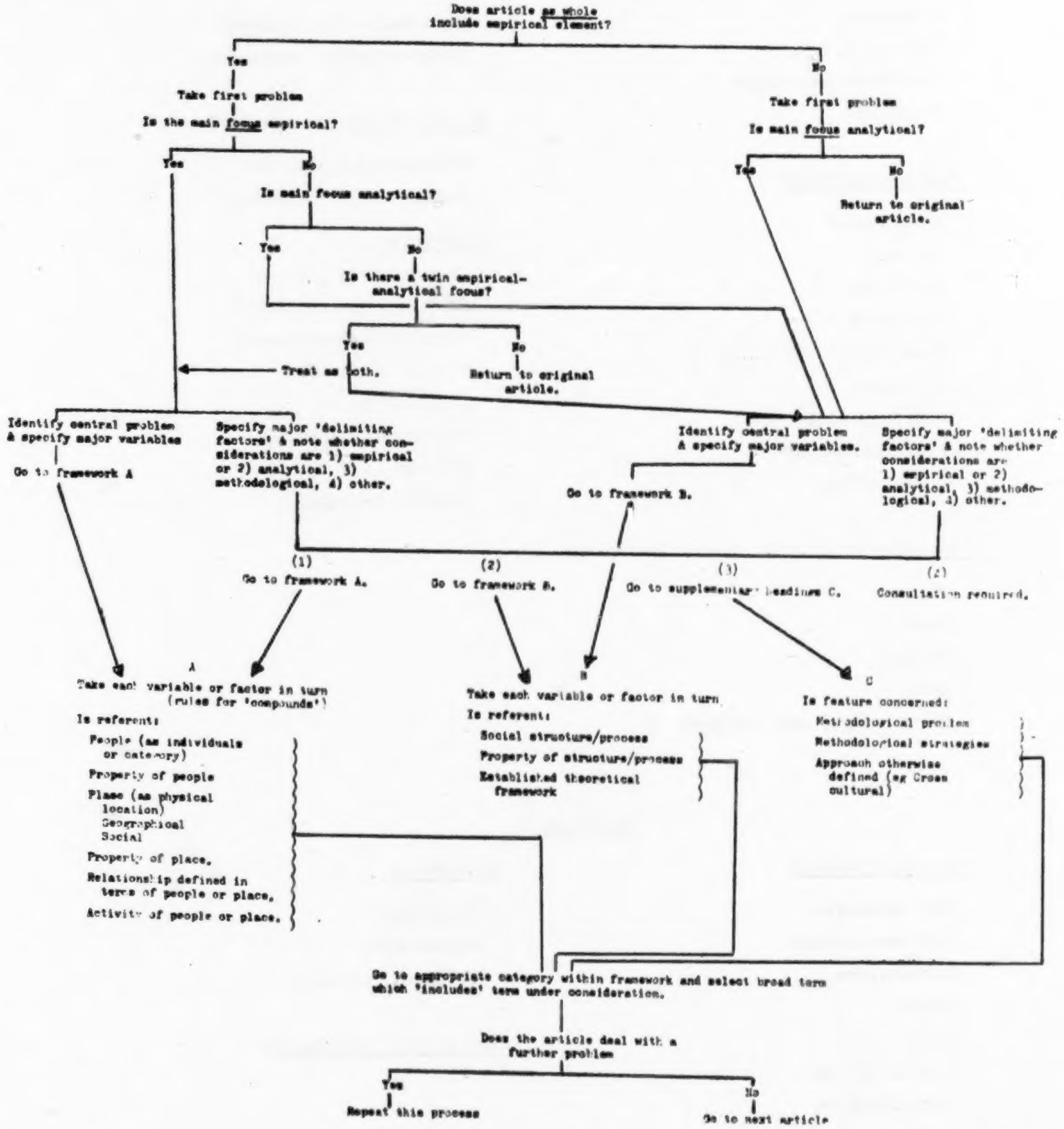
(e) methodological features require separate handling.

(See attached appendices for a possible implementation of these ideas.)

VW/RAC
12.1.72.

Appendix A to
Paper III 6A (iv)

SUBJECT HEADINGS: DECISION-MAKING GUIDE



Proposed subject headings

Framework A

People (individuals, categories)

- Academics
- Fisher, H.A.L.
- Religious categories
- Teachers

People properties

- Achievement
- Anxiety
- Attitudes
- Behaviour
- Demographic variables
- Language
- Personal variables
- Professionalism
- Social class

Place (geographical, social locations)

- Germany
- Home
- India
- School
- Universities and colleges

Place properties

- Curriculum (ie courses)
- Socio-economic conditions

Relationships (people defined)

- Friendship relations
- Parent-child relations

Activities

- Community activities
- Compensatory education
- Curriculum development
- Higher education
- Legislation
- Policy and planning
- Protest
- Teacher training

Framework B

Structure/process

- Bureaucracies
- Differentiation
- Disciplines
- Groups
- Role
- Social change
- Socialisation

Properties

- Knowledge
- Opportunity
- Political ideology

Theoretical frameworks

C. Methodology

- Historical studies
- Methodological problems

? **D. Other**

- Costs and benefits
- Equality

EXAMPLE 1

Use of a Parsonian framework involving consideration of functional prerequisites of social systems (adaptive, goal attainment, pattern maintenance and integrative functions). For analysis of students' reasons for becoming teachers in terms of preference for different roles (teacher as worker, educator, person and teacher) associated with teaching, taking account of age, sex and academic qualifications. Source of data - Sample of 450 students at one college of education in Scotland.

Does article as whole include empirical element? YES

First problem (there is only one in this case).

Is main focus empirical? YES

Identify central problem and specify major variables.

Reasons for becoming teacher
→ Framework A

Specify major 'delimiting factors'.

Student teachers
→ Framework A

Role
→ Framework B

Parsonian framework
→ Framework B

Go to categories of Framework A.

Reasons for becoming teachers
= Property of people, or Activity?
Round up to broader term

Student teachers
= People
Round up to ?

Go to categories of framework B.

Role = Structure/Process
Round up further ?

Parsonian framework
= Theoretical framework
Round up further ?

EXAMPLE 2

Proposal of definition of goals in formal organisations, based on Parsonian definition in terms of functional imperatives of social systems, distinguishing between output and support goals. Illustrated by data on American universities.

Does article as whole include empirical element? YES

First problem (there is only one in this case).

Is main focus empirical? NO

Is main focus analytical? YES

Identify central problem and specify major variables.

Goals
→ Framework B

Formal organisations
→ Framework B

Specify major 'delimiting factors'.

Parsonian definition...
→ Framework B

Universities
→ Framework A

Go to categories of Framework A.

Universities
= Place
Round up further ?

Go to categories of Framework B.

Goals
= Property
Round up further?

Formal organisations
= Structure
Round up further ?

Parsonian definition...
= Theoretical framework
Round up further ?

Appendix B to
Paper III 6A (iv)

Sample index of selected items

This index, ordinary in appearance though lacking a full set of cross-references, was compiled to test particular procedures for assigning subject headings to documents. The index is to be viewed not in terms of adequacy or internal consistency of headings, but as a collective representation of what is significant in a set of documents. The procedures were designed so as to allow us, within the constraints of a conventional framework of subject headings, to index the maximum number of significant aspects of documents such as those contained in the sample. Inspection of the procedures will reveal that there are further aspects, at least as likely to influence retrieval of documents, which we are precluded from building into the index.

ACADEMICS

(Suggest subdivide by level of establishment.)

Exploration of notion that young college faculty members differ from more experienced colleagues in terms of role definitions: investigation of role preferences (preferred allocation of time amongst teaching, research, committee work and student counselling, and commitment to each role) and role performances (actual allocation of time). Source of data - Sample of 120 faculty members from 14 liberal arts colleges in US. 14

See also Teachers.

ACHIEVEMENT

(Suggest divide into occupational and educational achievement at main heading level. Suggest use formal subdivisions; subdivision by group studied an additional possibility.)

Investigation of personal and home background characteristics of able misfits (under-achievers with high IQ); bearing on measures to deal with their problems. Source of data - Sample of 103 school age children in Great Britain. 12

ANXIETY

(Suggest subdivision by group studied.)

Consideration of relations amongst anxiety about school, occupational class and coeducation, taking account of sex. Source of data - Sample of 1,120 children (age 11+) from 42 grammar schools (boys, girls and mixed) in England and Wales. 6

ATTITUDES

(Suggest no subdivision simply refer to specific variables: subdivision by group studied an additional possibility if required.)

Model of relationship between verbal attitude and behaviour towards object of attitude: social constraint and social distance as intervening (situational) variables. Tested in investigation of prejudice towards negroes amongst college students. Source of data - Sample of 537 freshmen in one university in U.S. 17

BEHAVIOUR

(Suggest subdivision by group studied.)

Model of relationship between verbal attitude and behaviour towards object of attitude: social constraint and social distance as intervening (situational) variables. Tested in investigation of prejudice towards negroes amongst college students. Source of data - Sample of 537 freshmen in one university in U.S. 17

BILINGUALISM

See Language.

BUREAUCRACIES

(Suggest subdivide by type or by process for which studied eg decision making.)

Rational principles governing administrative bureaucracies and lack of faculty involvement in attempt to change the decision-making system: explanation for sequence of events leading to Berkeley war. 16

COMMUNITY ACTIVITIES

(Suggest subdivide by group.)

Thesis that effectiveness of teachers in inner city community schools depends on a professionalism which involves not merely instruction but also playing a part in curriculum materials development and close involvement in community activities; model for training programme proposed and compared with traditional programmes; implications in terms of role conflict, and costs and benefits. 5

COMPENSATORY EDUCATION

(Suggest main heading might be Education subdivided by level of education and specific form eg compensatory education. Alternatively subdivide by specific establishment or country.)

Discussion of own theory (elaborated and restricted linguistic codes), emphasising that use of particularistic meanings at home does not mean that working class child is incapable of learning to use universalistic meanings in another (school) context under suitable conditions; critique of assumptions underlying notion of compensatory education. 3

COSTS AND BENEFITS

(Suggest subsume under main heading Economic considerations.)

Thesis that effectiveness of teachers in inner city community schools depends on a professionalism which involves not merely instruction but also playing a part in curriculum materials development and close involvement in community activities; model for training programme proposed and compared with traditional programmes; implications in terms of role conflict, and costs and benefits. 5

CURRICULUM (ie COURSES)

(Suggest subdivide by subjects of the curriculum.)

Examination of social science courses in engineering education in advanced societies (Britain, Federal Republic of Germany), with regard to relevance to conditions in developing societies (India). 13

CURRICULUM DEVELOPMENT

(Suggest subsume under main heading Educational Research, subdivided by research area.)

Thesis that effectiveness of teachers in inner city community schools depends on a professionalism which involves not merely instruction but also playing a part in curriculum materials development and close involvement in community activities; model for training programme proposed and compared with traditional programmes; implications in terms of role conflict, and costs and benefits. 5

DEMOGRAPHIC VARIABLES

(Suggest no subdivision simply refer to specific variables; subdivision by group studied an additional possibility if required.)

Consideration of relevance of a sociolinguistic theory to study of bilingualism, emphasising societal role (institutional contexts and functional aspects) of mother/other tongue. Background to investigation of relations amongst function (understanding/reading/writing/speaking), developmental phase, context (home/work/religion) and demographic variables in a stable bilingual community; attempt to assess reliability and meaningfulness of sociolinguistic data collected by self report census methods. Source of data - Sample of 90 Puerto Rican households in U.S. 10

DIFFERENTIATION

(Suggest divide at main heading level into educational/non-educational and subdivide by specific differentiating factors eg race.)

Discussion of utility of notion of differentiation as analytic tool for study of dynamics of educational change and of educational change in context of social change. Implications for educational planning. 1

DISCIPLINES

(Suggest subdivide by specific disciplines.)

- Analysis (in historical perspective) of factors in present crisis facing universities with regard to the legitimacy of the pursuit of knowledge as such (tracing change in attitudes to scientific enquiry - from C17 confidence that knowledge ultimately enhances human power to influence society, through C19 retreat from 'reason' to subjectivism, to current ethic of social despair impairing belief in possibility of progress through knowledge and demanding an involvement on the part of universities which would destroy the detachment essential to reasoned enquiry); implications for the discipline of sociology. 2

EQUALITY

(Suggest subsume under main heading Philosophical considerations.)

Consideration of ideas about the pool of ability, social justice and economic wastage. Background to discussion of alternative definitions of notion of equality of educational opportunity (move from standardised education for all regardless of ability to standardised education for those of particular ability regardless of home background); proposal for new definition (positive discrimination for educationally underprivileged). 9

FISHER, H.A.L.

(Subdivision under names of individuals seems inappropriate.)

Outline of Fisher's contribution to educational reform (1914-22), viewed against the background of the political and economic situation, and with particular reference to the events leading to the implementation of the 1918 Act. 7

FRIENDSHIP RELATIONS

(Suggest subsume under a main heading (Social relations) subdivided by specific people-defined relations.)

Study of alternative models (emphasising differences between/within groups) of structure of social relationships in religious subcommunities. Based on analysis of friendship relationships among religious and ethno-religious groups as determined by broad category of preference within which given denomination is found, group's socio-economic position and characteristics of its religious beliefs and organisations. Source of data - Sample of 1,013 native born males (ages 21 - 64) in U.S. 15

GERMANY (FEDERAL REPUBLIC)

(Suggest two main headings (Advanced countries and Developing countries) to subsume all individual countries. Subdivision by individual country names.)

Examination of social science courses in engineering education in advanced societies (Britain, Federal Republic of Germany), with regard to relevance to conditions in developing societies (India). 13

GROUPS

(Suggest divide at main heading level into Groups (ie experimental groups and Groups (ie socially defined groups) and subdivide former by specific process eg decision making and latter by specific group delimiters eg ref. groups.)

Study of alternative models (emphasising differences between/within groups) of structure of social relationships in religious subcommunities. Based on analysis of friendship relationships among religious and ethno-religious groups as determined by broad category of preference within which given denomination is found, group's socio-economic position and characteristics of its religious beliefs and organisations. Source of data - Sample of 1,013 native born males (ages 21 - 64) in U.S. 15

HIGHER EDUCATION

(Suggest main heading might be Education subdivided by level of education and specific form eg compensatory education. Alternatively subdivide by specific establishment or country.)

Examination of social science courses in engineering education in advanced societies (Britain, Federal Republic of Germany), with regard to relevance to conditions in developing societies (India). 13

Discussion of recent developments in higher education in India, in context of social problems of an emerging nation and specifically Indian problems. Implications for education in other countries. 18

See also Universities and colleges.

HISTORICAL STUDIES

(Suggest chronological subdivision if required.)

Analysis (in historical perspective) of factors in present crisis facing universities with regard to the legitimacy of the pursuit of knowledge as such (tracing change in attitudes to scientific enquiry - from C17 confidence that knowledge ultimately enhances human power to influence society, through C19 retreat from 'reason' to subjectivism, to current ethic of social despair impairing belief in possibility of progress through knowledge and demanding an involvement on the part of universities which would destroy the detachment essential to reasoned enquiry); implications for the discipline of sociology. 2

Outline of Fisher's contribution to educational reform (1914-22), viewed against the background of the political and economic situation, and with particular reference to the events leading to the implementation of the 1918 Act. 7

HOME

(Suggest subsuming under a main heading (Social situations) and subdividing by the specific situation. Some items might be more appropriately located under Home background variables.)

Discussion of own theory (elaborated and restricted linguistic codes), emphasising that use of particularistic meanings at home does not mean that working class child is incapable of learning to use universalistic meanings in another (school) context under suitable conditions; critique of assumptions underlying notion of compensatory education. 3

Presentation of model of social learning which links differential emphasis on use of language, different areas of orientation (basic skills/interpersonal relations) and different forms of social relations within the social structure, reflecting different implicit theories of learning (self regulating/didactic) which affect child's concept of self and of role relationships. Leading from investigation of social class differences in perception of relevance of language in familial socialisation, distinguishing between interpersonal socialisation and socialisation into basic skills. Implications for sources of discontinuity for the working class child between home and school. Source of data - Sample of 100 mothers in England. 4

IDEOLOGY

See Political ideology

INDIA

(Suggest two main headings (Advanced countries and Developing countries) to subsume all individual countries. Subdivision by individual country names.)

Examination of social science courses in engineering education in advanced societies (Britain, Federal Republic of Germany), with regard to relevance to conditions in developing societies (India). 13

Discussion of recent developments in higher education in India, in context of social problems of an emerging nation and specifically Indian problems. Implications for education in other countries. 18

KNOWLEDGE

(Suggest subdivide on individual/institutional basis.)

Analysis (in historical perspective) of factors in present crisis facing universities with regard to the legitimacy of the pursuit of knowledge as such (tracing change in attitudes to scientific enquiry - from C17 confidence that knowledge ultimately enhances human power to influence society, through C19 retreat from 'reason' to subjectivism, to current ethic of social despair impairing belief in possibility of progress through knowledge and demanding an involvement on the part of universities which would destroy the detachment essential to reasoned enquiry); implications for the discipline of sociology. 2

LANGUAGE

(Suggest use formal subdivisions; subdivision by group studied an additional possibility.)

Discussion of own theory (elaborated and restricted linguistic codes), emphasising that use of particularistic meanings at home does not mean that working class child is incapable of learning to use universalistic meanings in another (school) context under suitable conditions; critique of assumptions underlying notion of compensatory education. 3

Presentation of model of social learning which links differential emphasis on use of language, different areas of orientation (basic skills/interpersonal relations) and different forms of social relations within the social structure reflecting different implicit theories of learning (self regulating/didactic) which affect child's concept of self and of role relationships. Leading from investigation of social class differences in perception of relevance of language in familial socialisation, distinguishing between interpersonal socialisation and socialisation into basic skills. Implications for sources of discontinuity for the working class child between home and school. Source of data - Sample of 100 mothers in England. 4

Consideration of relevance of a sociolinguistic theory to study of bilingualism, emphasising societal role (institutional contexts and functional aspects) of mother/other tongue. Background to investigation of relations amongst function (understanding/reading/writing/speaking), developmental phase, context (home/work/religion) and demographic variables in a stable bilingual community; attempt to assess reliability and meaningfulness of sociolinguistic data collected by self report census methods. Source of data - Sample of 90 Puerto Rican households in U.S. 10

LEGISLATION

(Suggest divide at main heading level by educational/non-educational and subdivide chronologically if required.)

Outline of Fisher's contribution to educational reform (1914-22), viewed against the background of the political and economic situation, and with particular reference to the events leading to the implementation of the 1918 Act. 7

METHODOLOGICAL PROBLEMS

(Suggest subdivide by research area or by specific methodological problem dealt with.)

Consideration of relevance of a sociolinguistic theory to study of bilingualism, emphasising societal role (institutional contexts and functional aspects) of mother/other tongue. Background to investigation of relations amongst function (understanding/reading/writing/speaking), developmental phase, context (home/work/religion) and demographic variables in a stable bilingual community; attempt to assess reliability and meaningfulness of sociolinguistic data collected by self report census methods. Source of data - Sample of 90 Puerto Rican households in U.S. 10

OPPORTUNITY

(Suggest subdivide on individual/institutional basis.)

Consideration of ideas about the pool of ability, social justice and economic wastage. Background to discussion of alternative definitions of notion of equality of educational opportunity (move from standardised education for all regardless of ability to standardised education for those of particular ability regardless of home background); proposal for new definition (positive discrimination for educationally underprivileged). 9

PARENT-CHILD RELATIONS

(Suggest subsume under a main heading (Social relations) subdivided by specific people-defined relations.)

Presentation of model of social learning which links differential emphasis on use of language, different areas of orientation (basic skills/interpersonal relations) and different forms of social relations within the social structure, reflecting different implicit theories of learning (self regulating/didactic) which affect child's concept of self and of role relationships. Leading from investigation of social class differences in perception of relevance of language in familial socialisation, distinguishing between interpersonal socialisation and socialisation into basic skills. Implications for sources of discontinuity for the working class child between home and school. Source of data - Sample of 100 mothers in England. 4

PERSONAL VARIABLES

(Suggest no subdivision simply refer to specific variables: subdivision by group studied an additional possibility if required.)

Investigation of personal and home background characteristics of able misfits (under-achievers with high IQ); bearing on measures to deal with their problems. Source of data - Sample of 103 school age children in Great Britain. 12

POLICY AND PLANNING

(Suggest divide at main heading level by educational/non-educational and subdivide by specific planning problem.)

Discussion of utility of notion of differentiation as analytic tool for study of dynamics of educational change and of educational change in context of social change. Implications for educational planning. 1

POLITICAL IDEOLOGY

(Suggest subdivide on individual/institutional basis.)

Thesis concerning anarchist nature of contemporary radicalism (seen as central element in American culture, and as associated with decline in teleological sense resulting from religious crisis); implications for measures to deal with student unrest. 11

PROFESSIONALISM

(Suggest subdivision by group studied.)

Thesis that effectiveness of teachers in inner city community schools depends on a professionalism which involves not merely instruction but also playing a part in curriculum materials development and close involvement in community activities; model for training programme proposed and compared with traditional programmes; implications in terms of role conflict, and costs and benefits. 5

PROTEST

(Suggest divide at main heading level by educational/non-educational and subdivide further by group or setting if appropriate.)

Thesis concerning anarchist nature of contemporary radicalism (seen as central element in American culture, and as associated with decline in teleological sense resulting from religious crisis); implications for measures to deal with student unrest. 11

Rational principles governing administrative bureaucracies and lack of faculty involvement in attempt to change the decision-making system: explanation for sequence of events leading to Berkeley war. 16

As sociological notion, see Conflict.

RELIGIOUS CATEGORIES

(Subdivision might be by denomination if required.)

Study of alternative models (emphasising differences between/within groups) of structure of social relationships in religious subcommunities. Based on analysis of friendship relationships among religious and ethno-religious groups as determined by broad category of preference within which given denomination is found, group's socio-economic position and characteristics of its religious beliefs and organisations. Source of data - Sample of 1,013 native born males (ages 21 - 64) in U.S. 15

ROLE

(Suggest subdivide by specific roles (people) and by role processes eg role conflict.)

Consideration of problems of convergence and overlap, and of distinguishing between behaviour and expectations, in analysis of role; viewed as essential preliminary to development of model of teacher role with specific characteristics ((a) should specify behaviours and expectations unique to teachers and shared with other groups, (b) should take account of situations, (c) would require additional specific concepts). 19

Thesis that effectiveness of teachers in inner city community schools depends on a professionalism which involves not merely instruction but also playing a part in curriculum materials development and close involvement in community activities; model for training programme proposed and compared with traditional programmes; implications in terms of role conflict, and costs and benefits. 5

Exploration of notion that young college faculty members differ from more experienced colleagues in terms of role definitions: investigation of role preferences (preferred allocation of time amongst teaching, research, committee work and student counselling, and commitment to each role) and role performances (actual allocation of time). Source of data - Sample of 120 faculty members from 14 liberal arts colleges in U.S. 14

SCHOOL

(Suggest subsuming under a main heading (Social situations) and subdividing by the specific situation. Some items might be more appropriately located under School background variables.)

Discussion of own theory (elaborated and restricted linguistic codes), emphasising that use of particularistic meanings at home does not mean that working class child is incapable of learning to use universalistic meanings in another (school) context under suitable conditions; critique of assumptions underlying notion of compensatory education. 3

SCHOOL Cont'd

Presentation of model of social learning which links differential emphasis on use of language, different areas of orientation (basic skills/interpersonal relations) and different forms of social relations within the social structure, reflecting different implicit theories of learning (self regulating/didactic) which affect child's concept of self and of role relationships. Leading from investigation of social class differences in perception of relevance of language in familial socialisation, distinguishing between interpersonal socialisation and socialisation into basic skills. Implications for sources of discontinuity for the working class child between home and school. Source of data - Sample of 100 mothers in England. 4

Consideration of relations amongst anxiety about school, occupational class and coeducation, taking account of sex. Source of data - Sample of 1,120 children (age 11+) from 42 grammar schools (boys, girls and mixed) in England and Wales. 6

SOCIAL CHANGE

(Suggest divide at main heading level by educational/non-educational and subdivide further by group or setting if appropriate.)

Discussion of utility of notion of differentiation as analytic tool for study of dynamics of educational change and of educational change in context of social change. Implications for educational planning. 1

See also Legislation.

SOCIAL CLASS

(Suggest use formal subdivisions; subdivision by group studied an additional possibility.)

Discussion of own theory (elaborated and restricted linguistic codes), emphasising that use of particularistic meanings at home does not mean that working class child is incapable of learning to use universalistic meanings in another (school) context under suitable conditions; critique of assumptions underlying notion of compensatory education. 3

Presentation of model of social learning which links differential emphasis on use of language, different areas of orientation (basic skills / interpersonal relations) and different forms of social relations within the social structure, reflecting different implicit theories of learning (self regulating/didactic) which affect child's concept of self and of role relationships. Leading from investigation of social class differences in perception of relevance of language in familial socialisation, distinguishing between interpersonal socialisation and socialisation into basic skills. Implications for sources of discontinuity for the working class child between home and school. Source of data - Sample of 100 mothers in England. 4

Consideration of relations amongst anxiety about school, occupational class and coeducation, taking account of sex. Source of data - Sample of 1,120 children (age 11+) from 42 grammar schools (boys, girls and mixed) in England and Wales. 6

SOCIAL RELATIONS

See Friendship relations.

SOCIALISATION

(Suggest subdivide by group studied or by formal method.)

Presentation of model of social learning which links differential emphasis on use of language, different areas of orientation (basic skills/interpersonal relations) and different forms of social relations within the social structure, reflecting different implicit theories of learning (self regulating/didactic) which affect child's concept of self and of role relationships. Leading from investigation of social class differences in perception of relevance of language in familial socialisation, distinguishing between interpersonal socialisation and socialisation into basic skills. Implications for sources of discontinuity for the working class child between home and school. Source of data - Sample of 100 mothers in England. 4

SOCIO-ECONOMIC CONDITIONS

(Suggest subdivide by country.)

Examination of social science courses in engineering education in advanced societies (Britain, Federal Republic of Germany), with regard to relevance to conditions in developing societies (India). 13

Discussion of recent developments in higher education in India, in context of social problems of an emerging nation and specifically Indian problems. Implications for education in other countries. 18

TEACHER TRAINING

(Suggest subdivision by establishment. Alternatively main heading might be Professional training divided into educational/non-educational at main heading level and subdivided by specific group.)

Thesis that effectiveness of teachers in inner city community schools depends on a professionalism which involves not merely instruction but also playing a part in curriculum materials development and close involvement in community activities; model for training programme proposed and compared with traditional programmes; implications in terms of role conflict, and costs and benefits. 5

TEACHERS

(Suggest transfer to be subsumed under Academics.)

Consideration of problems of convergence and overlap, and of distinguishing between behaviour and expectations, in analysis of role; viewed as essential preliminary to development of model of teacher role with specific characteristics ((a) should specify behaviours and expectations unique to teachers and shared with other groups, (b) should take account of situations, (c) would require additional specific concepts). 19

TEACHERS Cont'd

Thesis that effectiveness of teachers in inner city community schools depends on a professionalism which involves not merely instruction but also playing a part in curriculum materials development and close involvement in community activities; model for training programme proposed and compared with traditional programmes; implications in terms of role conflict, and costs and benefits. 5

See also Academics.

THEORETICAL FRAMEWORKS

(Suggest subdivide by type of framework.)

Presentation of model of social learning which links differential emphasis on use of language, different areas of orientation (basic skills/interpersonal relations) and different forms of social relations within the social structure, reflecting different implicit theories of learning (self regulating/didactic) which affect child's concept of self and of role relationships. Leading from investigation of social class differences in perception of relevance of language in familial socialisation, distinguishing between interpersonal socialisation and socialisation into basic skills. Implications for sources of discontinuity for the working class child between home and school. Source of data - Sample of 100 mothers in England. 4

Proposal of synthesis of symbolic interactionism and ethnomethodology for joint analysis of social psychological and sociological problems. 8

Consideration of relevance of a sociolinguistic theory to study of bilingualism, emphasising societal role (institutional contexts and functional aspects) of mother/other tongue. Background to investigation of relations amongst function (understanding/reading/writing/speaking), developmental phase, context (home/work/religion) and demographic variables in a stable bilingual community; attempt to assess reliability and meaningfulness of sociolinguistic data collected by self report census methods. Source of data - Sample of 90 Puerto Rican households in U.S. 10

UNIVERSITIES AND COLLEGES

(Suggest subsuming under a main heading (Social situations) and subdividing by the specific situation. Some items might be more appropriately located under University (including college) background variables.)

Analysis (in historical perspective) of factors in present crisis facing universities with regard to the legitimacy of the pursuit of knowledge as such (tracing change in attitudes to scientific enquiry - from C17 confidence that knowledge ultimately enhances human power to influence society, through C19 retreat from 'reason' to subjectivism, to current ethic of social despair impairing belief in possibility of progress through knowledge and demanding an involvement on the part of universities which would destroy the detachment essential to reasoned enquiry); implications for the discipline of sociology. 2

UNIVERSITIES AND COLLEGES Cont'd

Rational principles governing administrative bureaucracies and lack of faculty involvement in attempt to change the decision-making system; explanation for sequence of events leading to Berkeley war. 16

See also Higher education.

Details of items to which sample index relates

1. Adams, Don and Farrell, Joseph P.
   Comparative Education, 1969, 5(3).
   - A. Policy and planning (= educational planning)
   - B. Differentiation
     Social change (= educational change)

2. Bendix, Reinhard
   American Sociological Review, 1970, 35(5).
   - A. Universities and colleges
   - B. Disciplines (= sociology)
     Knowledge
   - C. Historical studies

3. Bernstein, Basil.
   New Society, 1970, 387.
   - A. Compensatory education
     Home
     Language (= linguistic codes)
     School
     Social class (= working class)

4. Bernstein, Basil and Henderson, Dorothy.
   Sociology, 1969, 3(1).
   - A. Home
     Language
     Parent-child relations (= familial ....)
     School
     Social class
   - B. Socialisation
     Theoretical frameworks (= theories of learning)

5. Cuban, Larry
   Educational Research, 1969, 39(2).
   - A. Community activities
     Curriculum development
     Professionalism
     Teacher training (= training programmes)
     Teachers
   - B. Role
   - D. Costs and benefits

6. Dale, R.R.
   British Journal of Educational Psychology, 1969, 39(1).
   - A. Anxiety
     School (= grammar school)
     (= coeducation)
     Social class (= occupational class)

7. Dean, D.W.
   British Journal of Educational Studies, 1970, 18(3).
   - A. Fisher, H.A.L.
     Legislation (= educational reform)
   - C. Historical studies

8. Denzin, Norman K.
American Sociological Review, 1969, 34(6).

B. Theoretical frameworks (= symbolic interactionism)
(= ethnomethodology)

9. Evetts, Julia
British Journal of Sociology, 1970, 21(4).

B. Opportunity

D. Equality

10. Fishman, Joshua A.
American Journal of Sociology, 1969, 75(3).

A. Demographic variables
Language (= function)
(= developmental phase)
(= context)
(= bilingualism)

B. Theoretical frameworks (= sociolinguistic theory)

C. Methodological problems (= self report census methods)

11. Frye, Northrop.
New Society, 1969, 372.

A. Protest (= student unrest)

B. Political ideology (= anarchist nature of contemporary radicalism)

12. Pringle, M.L. Kellmer.
New Society, 1970, 410.

A. Achievement (= able misfits – underachievers with high IQ)
Personal variables (= personal and home background characteristics)

13. King, Anthony.
Comparative Education, 1969, 5(3).

A. Curriculum (ie courses) (= social science courses)
Germany
Higher education
India
Socio-economic conditions (= conditions in developing countries)

14. Klapper, Hope Lunin
Sociology of Education, 1969, 42(1).

A. Academics (= college faculty members)

B. Role

15. Laumann, Edward O.
American Sociological Review, 1969, 34(2).

A. Friendship relations
Religious categories (= religious and ethno-religious groups)

B. Groups

16. Marris, Peter.
    New Society, 1969, 354.

    A. Protest (= ... war)
       Universities and colleges (= Berkeley)

    B. Bureaucracies

17. Warner, Lyle G. and De Fleur, Melvin L.
    American Sociological Review, 1969, 34(2).

    A. Attitudes
       Behaviour

18. Wragg, Marie.
    Comparative Education, 1969, 5(2).

    A. Higher education
       India
       Socio-economic conditions (= social problems of an emerging nation)

19. Adams, Raymond S.
    Educational Research, 1970, 12(2).

    A. Teachers

    B. Role

III 6A (v) Report on meeting of Working Party held at Birkbeck College on 19th January 1972

The whole meeting was devoted to consideration of subject headings, the general problem being to structure the headings according to principles which will lead to appropriate and consistent organisation of material.

There was an initial discussion to clarify whether the concern was to integrate sociology and education or simply to combine them in the same index. It was agreed that there must be a 'dual view'.

This was followed by a brief recapitulation of the main points emerging from discussion and experiment to date. The major distinction between alternative viewpoints appeared to be a concrete/abstract distinction. This was a dimension other than and additional to the broad/related-or-coordinate/narrow dimension on which subject headings may be characterised. The researchers had experimented with frameworks of subject headings defined in this way, but the resulting indexing had been by any standards often inadequate, inappropriate or both. Classification of a term into one framework or another, and 'rounding up' to an appropriate broad term, had been guided by consideration of nature of term as defined by context. By working with more limited contextual units, more satisfactory results were obtained. A decision making chart had been prepared to sort material into one of the agreed frameworks or another on the basis of successive questions as to different aspects of context. As a further step tentative suggestions were included for 'rounding up' any given notion within its framework depending on nature of referent. Thus, working on the basis of the previously agreed broad distinction between frameworks, and the assumed need for differential distribution of material in the index, the object of the chart was to enable consistent handling of like material. The researchers asked for comment on the method and for assistance in more precise or appropriate formulation of the questions the indexer must ask in order to sort material in this way.

The Working Party recognised that the basic distinction might be operationally defined in various ways:

(a) Empirical/Theoretical variables.
(b) Thing being studied/Property of thing.
(c) Commonly used/Technical terms.
(d) Units/Structure-process.
(e) Concrete/Abstract.
(f) Individual/Institutional.
(g) Sociological/Educational.

It was agreed that several of these definitions might be equally appropriate to any given type of material and that no single definition could be applied to all types of material. The point was made that there was a serious risk of inconsistency with any scheme in which indexers were required to shift frequently from one kind of thinking to another. The further question was

Cont'd....

raised of whether the indexer should concern himself with referents of terms in some sense 'absolutely' conceived even if variously defined, or with terms as defined by authors, major definitional distinctions to be taken as indicating, for our practical purposes, different 'things'. There was no agreement on this point.

The suggestion was made that there might be a need for multiple rather than dual frameworks, but the problem of clarifying the defining principle(s) still remained. It was also pointed out that not only terms but also the concepts to which terms referred might be used in so many ways that there was the likelihood, however frameworks were defined, of ending up with almost the same terms in each framework. Further, there was felt to be an undesirable circularity in that the content of frameworks was being specified on the basis of questions associated with preconceptions about what this content should be, rather than on the basis of questions emerging from individual documents as handled (although there was recognition that the approach had emerged after experimentation by the researchers with many alternative procedures in the course of analysing and indexing about 200 documents).

As an alternative it was suggested that the major variables in each problem might simply be classified into one of the following categories:

A. People and people properties (individual/aggregate level).
B. Places and place properties (physical location).
C. Social structures/processes and their properties, including
   social relationships.
D. Culture and value systems.
E. Established theoretical frameworks.

Both sociologists and non-sociologists found these categories meaningful, but it was recognised that respective expectations of the scope and content of these categories would differ. Thus the main problem which the dual framework approach was proposed to solve remains - namely that of differential 'rounding up' to broad terms as appropriate to the document (as contrasted with simple word matching), and hence of differential distribution and accessibility of documents (cf coeducation example) in relation to alternative viewpoints upon education as an area of study. (The problem in relying upon subheadings for alternative modes of access, as argued in a paper circulated earlier, was noted.) The effect of such an approach would be to cater for either the sociologist or the non-sociologist at the expense of the other.

It seemed that if the dual approach could not be operationalised effectively, the only other alternatives were (a) to ch[illegible] between a broadly sociological or an educational operationalisation of such a scheme, (b) to index each document twice, or (c) to work on the basis of word matching, with the onus of creating meaningful sets of the documents thus co-located relegated to the subheading level (at which point the same basic problem of clarifying the principles to be adopted must again be faced if effective differential access in relation to the 'dual view' agreed upon earlier is an objective).

V /RAC
25.1.72.

III 6A (vi) Establishment of subject headings: continued experiment with alternative principles

1. The following points are guiding our thinking:
   - (i) A single framework of subject headings is inadequate to provide immediate access to material grouped appropriately for different groups of users. This is not merely a question of the terms used as access points, but has also to do with the expectations of the range of material grouped under any heading.
   - (ii) There is the further question of the validity of headings in relation to different types of work.
   - (iii) The principles adopted must take account of these points. It should be noted that our concern is not to index a given collection of documents, but to develop a system of indexing.

2. At the last meeting of the Working Party an empirical operationalisation of the agreed concrete/abstract distinction was rejected at a basis for alternative frameworks of subject headings. It also seemed that the term-in-context principle might have to give way to individual treatment of each variable.

3. Alternative operationalisations of the concrete/abstract distinction were therefore considered in relation to our requirements, namely that educational terms should not be generalised to terms not meaningful to the educationalist, and that non-education-specific terms should not be 'rounded up' to higher order abstractions than are found in a given document.

4. There was also experimentation with different sized contextual units: whole subject, central research problem + delimiting factors, individual variable.

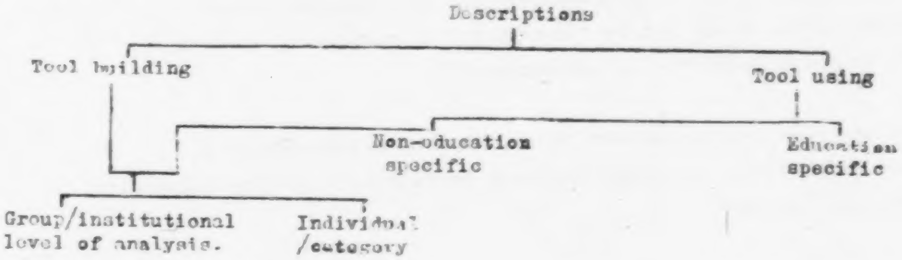
Cont'd....

5. With regard to feasibility of the various alternatives, a summary of the results is given in the following table. (Certain of the suggestions were excluded as leading to a simple framework situation; others were merged as having for practical purposes the same effect.)

Basis of framework	Size of contextual unit		
	Whole subject	Problem + delimiters	Individual variable
Empirical/ theoretical.	Produces split undesirable to sociologist.	Produces split undesirable to sociologist.	N/A
Context specific or not.	Both types of elements may occur.	Both types of elements may occur.	Possible
Level of analysis.	May be relationship between individual/institutional factors.	May be relationship between individual/institutional factors.	Worth exploring

6. Because of the form of the descriptions it is necessary to start from the empirical/theoretical elements of studies and, according to focus (eg tool building/tool using), to specify the problem in appropriate terms. From that point it seemed that handling individual variables in terms of context-specificity and 'level of analysis' might both be viable and sort material at least to our minimum requirements (section 3 above).

7. The following chart was developed and this combination of features tested:



8. This however proved equally unacceptable in that whilst, at a certain level, the resulting distribution of documents met the specified requirements, it was then seen to disperse material related in other important ways.

9. It was concluded that further consideration of requirements was essential.

VW/RAC 22.3.72.

III 6B (i) Report on Furzedown College meeting of Working Party 24/25th March 1972

General

Discussion is reported in terms of the rationale which emerged rather than strictly chronologically and is subject to confirmation by the Working Party as a record of the agreement reached.

This meeting in some sense marked a new phase of work in the project. Previous attempts at index construction had been based upon the notion of objective handling of material, since reproducibility and other tests of a mechanical kind had been built into the brief to which the researchers had to work. The search for principles appropriate to a system of this kind, yet at the same time meaningful to subject experts, was unsuccessful in that it produced an organisation of material unhelpful in relation to sociological and educational study.

The principal reason for this emerged in discussion of the content of the paper the Working Party are preparing for publication, namely that the patterning of knowledge is socially constructed. Thus the notion of system must be redefined to connote not the presentation of information within one or more structures or sets of categories, but its presentation in relation to the dimensions upon which users structure their knowledge. The aim should be to allow users to select relevant dimensions and themselves to decide upon and gather information into categories meaningful to the enquiry in hand.

This contrasts radically with the notion of 'translating' search problems into a set or sets of predetermined categories into which information is classified and stored by the system. There had been discussion with OSTI, and it had been agreed that there was a need to reconsider the nature of the task in which we were engaged, and to formulate detailed alternative specifications as a framework for a new phase of experimental work by the researchers along the lines suggested by the Working Party and for consideration by the Working Party.

Discussion of specifications centred around search strategy and index structure as interdependent problems.

Search strategy (ie use of index)

The term search strategy was found to be ambiguous and the heading of this section has therefore been modified. A distinction was made between purpose of search and operations or activities involved in searching, and it was felt that operations could be described in intellectual or 'technical' terms. It was agreed that technical considerations should be deferred until the intellectual aspects of searching had been further clarified.

In general the user would have a research problem which represented an information need (it was not accepted that the former would be 'translated' into the latter by means of cognitive 'strategies'). The nature of the problem would predispose the user to a particular kind of information tool from the available range, subject to considerations of 'cost' of use relative to commitment to task. Faced with a given tool it would also determine how the user would wish to proceed in his search. Tools appropriate to problems are required.

Cont'd....

Various categorisations of purpose (arising from need) were suggested.

(a) A first analysis might be as follows:

1. Finding what is known to exist;
2. Finding what is available;
3. Seeing if anything is available;
4. Seeing what is available in related fields;
5. Seeing what conceptual tools are available.

(b) Any of these requirements might be associated with any stage of enquiry (stages in no fixed order and enquiry starting at any point):

1. Reconnaissance.
2. Exploration.
3. Data collection.
4. Verification
5. Overview.
6. Loop back to (1).

(c) Teaching and research requirements were considered not to differ in kind but rather in terms of cut-off point and degree of selectivity.

(d) Similarly a creative or inspirational approach and a systematic approach were argued to be equally valid and purposeful as approaches, and were seen to differ not in kind but rather in terms of diffuseness/specificity and predictability of requirements. The former was distinguished sharply from browsing and serendipity, by virtue of progressive formulation and reformulation of problem in the course of a search.

It seemed that all those purposes would express themselves in two basic operations in index use:

1. Attempts to locate oneself within a body or bodies of knowledge (ie to identify others saying the kind of thing that the individual himself is saying, which may readily be translated into his frame of reference, or in other words to establish the genealogy of one's ideas);

2. Attempts to locate others within a body of bodies of knowledge (ie to identify alternative conceptions of the kind of problem one is working on, or languages in which to talk about a problem, into which one would translate one's own ideas; also to collect other data, however conceptualised, to relate to one's own ideas).

(A contrast was made with Kuhn's concept of paradigms, and reference made to Piaget and McHugh.)

Conceptual structure

This implied that at least two separate kinds of organisation must be involved - an everyday world one, and one representing different intellectual 'traditions' or epistemologies.

The former (everyday world) organisation would relate to people, places and things, and take account of those users who held an objectivist view of knowledge. Problems in this area have already been studied in some detail. It was suggested that the latter (by 'tradition') would involve the kinds of judgments made in compiling a critical review of the literature and was undesirable in that it might be seen as intellectually dishonest unless

Cont'd....

carried out by a leading expert and prefaced by a supporting essay explicitly delineating traditions. An acceptable alternative however would be to accept authors' statements of affiliation to a 'tradition', unless these were contradicted by the work they drew upon and cited as the 'genealogy' for their ideas. It was noted that citations should not be used as positive indicators of affiliation to 'tradition' except with supporting evidence from the text of the article. Consultations with authors on this matter might be envisaged also.

This might, by comparison with an organicist approach (cf Roget's Thesaurus), run the same risk of myth-creation and reification as, say, hierarchical schemes of organisation. Against this it was noted thata a safeguard was the multiplicity of myths (though clearly there were practical limits). The Working Party suggested also that the scheme would be seen as evolving through time. Knowledge was in a sense dictated by 'fashions'. Thus, as epistemologies were discarded, so they would disappear from the scheme, and as new epistemologies appeared or old ones were given a new lease of life so they could be added or reinstated without disruption of the scheme.

Also related to the time factor is the point that theories have content and scope. Thus fashions lead to clustering of work in terms of problems studied. A scheme with categories formed by relating Dawes' order/control dichotomy to level of analysis was shown to accommodate many such clusters:

Level of analysis	'Tradition'	
	Order	Control
Interplay of institutional areas (system & school level)	Education, economy and society Functions of education Typologies of systems	Struggle for school Historical, political developments.
Internal workings of systems	Role theory Streaming and grouping	Micropolitics Weberian organisational analysis Classroom management
Individual careers (sociol. definition)	Input models Socialisation and differentiation Attitudes and attainment Deprivation and privilege	Process models / / / Discrimination subcultures
Education as belief system	? Durkheim and Mannheim	Knowledge and control

This was still felt to be too rigid, however. Certain work could not readily be located within such a scheme. A social-psychological perspective would be somewhat different. It represented reification in that a determinist/voluntarist split offered the two sides of the same 'slice of the cake' rather than a different way of 'slicing' it.

Cont'd....

The Working Party therefore felt that the dimensions in this scheme should be handled separately, and further dimensions added in. Thus, in addition to (1) the 'real world' perspective, one would require (2) a 'problems' perspective (eg heredity v environment, social class and educational opportunity). Further to this, and in relation to a 'traditions' perspective, a feature of the intellectual organisation of an index should be (3) a range of dimensions which people currently use in structuring knowledge, or assumptions they will bring to their use of an index. One or more such dimensions might be relevant in characterising the tradition to which an author considers himself to belong. The following provisional list was put forward:

1. Level of analysis.
2. Determinism(structural, cultural, individual, biological etc. brands)/ voluntarism.
3. System/function.
4. Great men eg 'Weberian tradition'.
5. Positivist/phenomenological.
6. Macro/micro.
7. Normative/interpretive (ie Reified/non-reified).
8. Theory/empiricism.
9. ?

Social description and psychometric approaches were also touched upon.

There was discussion as to alternative ways of identifying such basic dimensions and evaluating them for the purposes of index construction. Collection and analysis of sociology of education course outlines was proposed, but there was some doubt that it would be possible to induce from these the basic dimensions or assumptions underlying the structuring of courses. Textbooks might afford some of the insights lacking in course outlines, but the necessary analysis would be time consuming. Essentially this was a problem in the sociology of knowledge, ie the processes by which (empirically) bodies of knowledge become organised, understanding of which should guide index construction. Since the index in any case requires now to be seen as a research instrument to be used in testing assumptions about the social construction of knowledge and the relation with index use, it was suggested that the lack of detailed investigation of this kind need not detract too seriously from our work.

Relation between index use and structure of index.

The Working Party conceived of the knowledge represented in the index being presented as having at least three 'faces' or perspectives - a real world, a problem and a tradition 'face'. Users might elect to approach the index via one or more of these depending on purpose and whether they 'spoke the language' or were able to adopt the perspective. There could be no simple 'transliteration' of one perspective into another, nor any simple equation of documents with knowledge.

It would be assumed that the user would make quite clear distinctions in terms of the perspective in which he was interested, in terms of the choice of dimensions offered him within any perspective, and in terms of the area of immediate interest. Crude distinctions (at the level of 'isms') were felt to come from the individual's working stance and to be

Cont'd....

used to characterise the material around the area of his interests. Thus he would be enabled to locate either himself and his ideas or his problem within the body of knowledge represented by the index, and to structure what he finds in a way meaningful to him. In other words, the user will 'choose his own way through', responding subjectively to what he finds.

This would characterise the nature of the intellectual structure which should underlie our index. Physical presentation of index (eg perspective by perspective / alphabetic interfiling of sections of material) would be a question for later consideration, together with the need for further substructure(s). An attempt will be made to translate the outline 'map' we have formulated into operational terms, as a basis for discussion and further development, in the light of our conception of index use, at the next meeting.

VW/RAC
27.3.72.

III 6B (ii) Report on meeting of Working Party at Birkbeck College, 28th April, 1972

Members of the Working Party tested a proforma and a set of guidelines prepared as a first attempt to operationalise the 'faces of knowledge' model outlined by the group at the last Fursedown meeting. A small set of documents selected by the researchers for the questions they raised was used for the purpose. Discussion led from the alternative ways in which, within the proposed framework, members considered it appropriate to handle these documents, to the underlying problems involved.

Real World

Sample data, though from some points of view trivial in the index context, has concreteness, and hence may be helpful to some users. Population is of more significance to others, though it is not always clear what population a sample has been taken to represent, so that there may be a risk of misleading the user. It was considered that the index must be regarded as experimental in the sense that it should enable us to discover more of user modes of approach and expectations of an index. It was therefore suggested that the researchers should explore the feasibility of including population on the grounds that, if possible, both kinds of data should be represented. Degree of use, and of satisfaction in use, would then be allowed to determine eventual practice. A similar approach might apply to the indexing of 'places'. A 'temporal' index was not considered necessary.

Problems

'Problems' was felt to be a misleading term in that clustering of work around problems (variously defined) would only emerge on later inspection of a body of indexing at the level of, say, individual variables. No alternative 'generic' term was suggested. It was considered to cover both (1) 'subject' or 'unit of study' (i.e. what the author is centrally concerned with - this might be more than one thing, and there might be more than one level of concern), and (2) the variables or concepts which are employed. In the case of the latter it was agreed that generalised statements should form the basis of indexing, rather than operational definitions. It was agreed also that 'implications' should be indexed where there was textual evidence to support this.

The researchers raised the question of compound terms which might be viewed as containing a 'contextual' variable, such as 'teacher role', or the 'subject' analysed as 'interpersonal relations in the classroom'. These might each (1) be split into two 'concepts', though later perhaps recombined as a 'cluster' in a genuine 'problems' index, or (2) regarded as a single idea. Though a single idea to the author, the invariable use of such terms or phrases as bases for wider headings would produce many sets of one. On the other hand it was pointed out that one of the drawbacks with the PRECIS system was that it depended upon breaking down such ideas to produce entry points which provided any clustering at all (though in fact the influence of the qualifier upon filing virtually prohibited clustering in the case of our material). The suggestion was that where a substantial cluster of work was known to exist we might anticipate the later step towards a 'problems' index. In other cases, it would be open to the indexer to split up compound ideas, though some concern was felt about doing so automatically. The problem was complicated by the possible effect of 'rounding up' author's terms to broader index headings.

Traditions

Unit of analysis was found difficult to apply and it was suggested that this should be discarded. There was preliminary discussion of the indexing of 'theoretical explanation(s) or model(s)'.

Further general points

(1) It was suggested that the researchers experiment with identifying significant aspects of a document and deciding to which area they belonged, by contrast with working section by section through the proforma looking for prescribed kinds of data.

(2) There was some feeling that snap judgments about an individual document based on necessarily brief study of the document (as in carrying out this exercise), even by subject experts, might well turn out on more thorough examination to be judgments which would need to be changed. The method however was adequate as a basis for discussion of problems.

OSTI/SEA
28.4.72.

Proposed guidelines for indexing

REAL WORLD

People

The information to be recorded here concerns the sample, rather than the population, under study and might include data relevant to the following personal characteristics:-

(i) Ascribed social characteristics: age, sex, birth order, etc.

(ii) Affiliation to social/political/economic institutions or strata, eg marital status, political affiliation, class position, etc.

(iii) Affiliation to educational institutions: primary, grammar, infant, etc.

(iv) Personal performances or qualities: motivation, ability, achievement, personality dimensions, IQ.

(v) Bio-medical personal characteristics: height, physique, etc.

Places

To be recorded here are:-

(i) Geographical setting, at the level of country or combination of countries.

(ii) Physical setting, at the level of institution, eg schools and colleges by level of education, or industrial firms. This form of heading is only required if the sample is specified in those terms ie sample of 6 high schools as opposed to sample of 512 pupils from 6 high schools.

PROBLEMS

Variables and concepts by which the research problem is specified

Only major variables will be considered, not minor ones, eg not 'age' and 'sex' in 'taking account of age and sex'. These are to be stated at the level of individual variables, but at a general level, rather than as the individual author has operationalised them. For instance, in the following example, 'achievement' rather than 'type of school' should be selected:-

Investigation of influence of sex on attitude towards school, according to type of school attended (taken as a measure of achievement).
Source of data - sample of 384 students (ages 14 - 15) from 11 schools (6 grammar and 5 secondary modern) in [? England].

Contextual as well as measured personal/organisational variables should be included here, eg 'ability' as in 'relation of X to Y amongst different ability groups' together with established dimensions, eg authoritarian - democratic dimension.
(Combinations of variables as headings can be expected to emerge later.)

Cont'd....

General 'phenomenon(a)' studied

By this we do not mean the variable taken as explanandum. We refer to any broader or higher order concept or statement of the problem given by the author eg 'stratification' in the following example:-

'Thesis that social class of student determines teacher's initial expectations of student academic success; differential effect on teacher-student interaction: growth of mutually accepted classroom stratification system.
Source of data - sample of one class of black ghetto children (kindergarten through second grade) in U.S. Observational study.'

A particular problem arises with compound concepts. Is, for instance, 'teacher role' to be regarded as a single concept or should it be regarded as two, ie 'teacher' and 'role', (cf also black militancy and occupational values). If the problem is simply stated at the level of variables no entry will be required here.

'Implications'

This refers to the situation in which the study contributes to understanding of some concept or problem not considered explicitly or in detail, eg 'educational planning' in the following example:-

'Discussion of utility of notion of differentiation as analytic tool for study of dynamics of educational change and of educational change in context of social change.
Implications for educational planning'.

Inference concerning the areas for which a study has implications would not be valid unless a very good case could be drawn up to support it.

TRADITIONS

Unit of analysis

This should be specified at the level of the problem and not at the level of sample, if this is different, ie for a study of cross ethnic differences studied amongst samples of negro and white students the unit of analysis should be the group and not the individual. If the impact of one unit, eg the organisation, on another unit, eg the student, is studied, then both should be specified. Suggested categories by which to specify unit of analysis are:-

A
1. Individual
2. Small group
3. Group (including family, social class and ethnic groups)
4. Community
5. Organisation
6. Institution
7. Society

B Social relationships
1. Within a group, community, organisation or society.
2. Across groups, communities, organisations or societies.

Nature of theoretical explanation(s) or model(s)

We are in the area here of theoretical debate. In the majority of cases, explanations or models are of the 'middle range'. The author may put forward a preferred explanation or model or debate the validity of alternatives. (Is it possible that classes of explanations may emerge which can be characterised in more general terms, at the level of, say, positivist/phenomenological or order/control?)

Cont'd....

'Significant others'

We propose to include not only historically great men to whom the author feels indebted, but also current writers who an author notes as having influenced his thinking. We are noting name of writer together with the date of his contribution, but not the specific work in question. The basis for selecting any given person cited should be on the strength of the author's comments about the work in question and not simply by the fact of citation. Possibly only those persons whose work the author actually builds on (we exclude citation to his own work) should be considered, and not those whose work he dismisses or reacts against. The basis of the specific selection might be that the person cited provided the theoretical environment in which the author is to work, provided a research instrument which is tested or validated by the author (but not just simply used as it stands), or whose findings provided questions for further investigation. Citations in support of particular points in an argument, rather than relating to the study as a whole, would not be included. Until it is possible to specify more clearly the reasons for selecting any given citation the reason for selecting it will be recorded.

'LITERARY FORM'

Overall properties of the particular study should be specified here eg 'review of research', 'critique', 'historical study', 'cross-cultural study' or 'longitudinal study'. If thought to be appropriate such labels as 'theory building' might also be specified.

DW/RAC
25.4.72.

III 6B (iii) Intellectual organisation for material relevant to the sociological analysis of education: aid to literature searching, August 1972

Aim

To develop a scheme in which concise descriptions of documents of use in the sociological analysis of education[1] can be searched, and relevant documents identified, on the basis of the criteria which the subject expert would use in searching a collection of actual documents. Factors which may be taken into account include both conceptual ones ('tradition' in which a researcher is working) and technical ones (the variables, methods etc. employed in the operationalisation and empirical investigation of a given problem).

The ideal method of proceeding is cover-to-cover searching of originals, since relevance essentially depends upon personal interpretation of documents, but this is time-consuming. Abstracts[2] are helpful, and in particular, systematic ordering saves the searcher's time, but only if the principle(s) upon which material is ordered match(es) the criterion(a) of selection he wishes to employ. Searchers may be expected to vary their criteria in relation to their perspective. A unitary system (one based upon a single organising principle such as substantive area of the social world) will interfere with the search when the user wishes to work material on a different basis.

We seek, therefore, to develop a system in which a given body of material is presented in terms of a variety of classifications.[3] This will inevitably reduce the searcher's capability to tailor his search to his problem (in other words to devise a search which responds to all (and only those) criteria relevant to his problem) by comparison with the total freedom of cover-to-cover searching. At the same time, his approach need not be wholly predetermined if he has a range of alternative classifications to choose from, and may switch amongst them so as to approximate as closely as possible to the fully 'tailored' search. These classifications, and the way they are related together as a system, should offer him a choice of search strategies broadly appropriate to different definitions of his problem and modes of enquiry. In this way time may be saved in literature searching without too serious interference, at the intellectual level, in the process.

---

[1] See Appendix A. (Not included here)

[2] See Appendix B. (Not included here)

[3] See Appendix C. A technical point to be made is that the distinction being made is not that between alphabetical/classified ordering of a given set of categories or concepts, nor does it concern degree of specifity of headings representing given objects or concepts, or indeed any formal aspect of system construction or search strategy. The aim is rather to offer alternative conceptual approaches to a given body of material, recognising that there are different views of the social world. Hence we refer to alternative classifications. A widely used solution to the problems of document retrieval is the development of a retrieval language, by means of which the terms describing document subjects and searchers' problems are translated into a common ambiguity-free language. Once reformulated, problems are matched against subjects to identify relevant documents. This solution has had to be rejected because, in a social science area, not merely the terms used but the phenomena to which terms refer are problematic, and perspective-dependent. Translation into a common language at best scrambles important distinctions which are generally meaningful in their intellectual context. At worst it makes certain ideas unavailable altogether in so far as disambiguation implies use of a single perspective.

Anticipated modes of use

Subject experts have suggested two basic orientations which characterise literature searching:
(1) the researcher may wish to locate his already formulated ideas in a 'tradition', with the object perhaps of legitimating them, and possibly in a search for data or guidance in empirical investigation to test their validity;
(2) the researcher may start with an empirical question suggested by a body of data, and may wish to formulate his ideas, to find a 'tradition' which will afford appropriate and helpful insights into his problem.

In practice, the searcher probably alternates between modes (1) and (2). The process of relating of data to ideas, which in turn directs the further search for data, is perhaps in one sense artificially distinguished in terms of the proposed modes (1) and (2). Nevertheless, the actual relations perceived, and the direction of the actual search which is devised, will necessarily vary from one searcher to another. For this reason, in the context of an information service, it is important to make the distinction, and to leave open to the individual searcher at least a broad range of options in constructing search strategies, even if not total freedom of choice. The alternative, the imposition of a single structure of pre-established relationships amongst data and ideas, has been unacceptable to subject experts we have consulted.

Implications for system construction

To accommodate these modes of use, the system should offer two kinds of classification: (1) Classifications by 'tradition'; (2) Classification in terms of operational aspects of research.

The actual classifications to be included in each category are a matter for investigation by analysis of the literature and consultation with subject experts. At this stage, the policy is to search for distinctive ways in which subject experts are considered to define and organise their ideas about the social world, and manage their research into it. Material is being sorted experimentally in terms of each of this range of alternative dimensions. Decisions on retaining, merging or discarding a given classification, or dimension of a classification, are being made on subject experts' advice concerning (a) its distinctiveness from others (b) its use in selecting relevant documents.

As many ways of classifying 'traditions' may be employed within the scheme as are widely employed by subject experts. Most 'traditions' appear to unpack unto a combination of elements of beliefs, assumptions, etc. Using those as dimensions of classification, each classification should specify a set of perspectives. Each perspective may be expected to focus upon a particular range of research problems by virtue of its particular construction of the social world.

Consequently, the more detailed organisation of material associated with a given perspective should reflect the organisation of ideas afforded by that perspective. It would be expected that the organisation of material within a given perspective would be unique to it. Identical groupings across perspectives would guide the system designers to consider the possibilities of (a) distortion (b) non-distinctive perspectives. This too is a matter for investigation.

With regard to classification in terms of operational aspects of research, we are concerned here with the operational aspects of research rather than the differing natures of ideas which guide it. Hence the various locations a document receives in this range of classifications sum to a total description of the document. This contrasts with the 'traditions' range of classifications in which the locations in alternative classifications are incompatible characterisations of a given document as a whole.*

---

* In technical terms, the former is the technique known as 'faceting'. (We envisage post-coordinate use will be more acceptable to our users than, say, a pre-coordinated rotated presentation.) The latter is most closely analogous to separate author, title and subject catalogues, allowing alternative modes of access to the same body of documents. Again we believe it will be more helpful to present our classifications separately, rather than integrating the various subsets into a single sequence with appropriate cross references. Presentation will, however, be linked to considerations of search strategies, and understanding of the latter is as yet limited.

Possible elements of the system are as follows:

I General introduction on nature of scheme and ways in which it may be used.

II Outline of sections.

III Details of documents in the system, arranged by journal and date.

IV Classifications by 'tradition'.
 eg by 'isms'.
  by 'view of the world'/mode of explanation.
  by 'significant other'.

V Classifications in terms of operational aspects of research.
 eg analyses of variables.
  methods.
  samples.

The system will also include listings by author, institutional affiliation. (See Appendix C.)

Operation of the system

Documents being entered into the system may be seen as passing through a series of sorting frames. The document may be classified on as many dimensions as apply. Users might think of themselves as 'processing' their research problems in exactly the same way through the series of sorting frames to locate cells containing abstracts of work by authors whose perspective or operational approach is appropriate to their problem.

Ideally an abstract of a document should be provided at all the points in the series of classifications where a document is located. In practice, it may be necessary to refer users from the classifications to the journal listing of items for abstracts and bibliographical details. This may not be a serious inconvenience, and condensed bibliographical details (eg Bloggs, BJS, 1969, 21-30) may be given at locations in the classifications for those who do not wish to take a step intermediate to reference to original documents in consulting abstracts.

As with other considerations of presentation, decisions will be dependent upon gradual clarification of the nature of the search strategies which the system should support. This clarification will come from subject experts' comments upon successive 'realisations' of perspectives in the system which will enable the specification of differentially appropriate search strategies and the progressive firming up of the system. Such firming up will not rule out sensitivity to change and ongoing modification as perspectives change. The assistance of subject experts is essential to a scheme such as this which is intended to support the literature searching process without interfering in it.

Future work

Clearly, perspectives are not entities which exist as entities - they are realised in use. Hence one would not expect to construct a system which takes the notion of perspective as its basic organising idea purely on evidence from documents, though these data are relevant. Of equal importance are the perspectives which subject experts employ in their research and hence in searching the literature, and how these interact with those of authors in an ongoing search as a search strategy is formulated and reformulated.

For this reason we are seeking the help of subject experts and hope for on-going consultation with them. In the first instance, we are anxious for comments on the general conception of the system. Following this we shall focus on the problem of clarifying the nature of the search strategies of which the system should take account. As a preliminary step in this direction, we have appended some of our ideas for the system to enable subject experts to clarify their own.

Once the general framework has been agreed, it will be possible to rough-sort material into it (in consultation with authors). The next step might be that, with this as a browsing facility, subject experts would be asked to create search strategies around problems of their choice as a starting point for further discussion and clarification of conceptual strategies, and to identify characteristic features which might be built in to the system. The aim would be to construct a system which organises material in terms of these characteristic features in searching the literature.

It is recognised that this is to objectify what is perhaps ideally left intuitive and fluid. On the other hand the system would, broadly at least, reflect the range of thinking in the field (it would not attempt to rationalise), as well as being responsive to new thinking, and as such might be seen to represent a practical aid, albeit a rough and ready one, at the intellectual level.

VW/DW/RAC
August 1972.

Discussion and explication of system of intellectual organisation

Elements of the system

I General introduction on nature of scheme and ways in which it may be used.

II Outline of sections.

III Details of documents in the system, arranged by journal and date.

IV Classifications by 'tradition'.
  eg by 'isms'.
    by 'view of the world'/mode of explanation.
    by 'significant other'.

V Classifications in terms of operational aspects of research.
  eg analyses of variables.
    methods.
    samples.

The system will also include listings by author, institutional affiliation.

In explaining in more detail the nature of these elements, in particular Sections IV and V, we make no claim to exhaustivity or finality. These are essentially ideas towards the system which subject experts are helping us to construct. We are casting our net widely to start with so as to identify all the options, on the principle of eliminating what is redundant later. Important options overlooked now are not likely to be raised if we exclude them from consideration by virtue of preconceptions built into the system. The content of this paper is for discussion first with experts whose schemas we refer to, then with authors for their comments on proposed placings of their work, as well as other subject experts.

I General introduction on nature of scheme and ways in which it may be used. The description of the system will be broadly similar to the foregoing paper.

II Outline of sections. The outline of contents should give a broad overview in sufficient detail to enable users to identify appropriate perspectives or operational aspects of their problem and should refer them to the relevant pages.

III Details of documents in the system, arranged by journal and date. Details of items in the system would be presented in a form similar to that illustrated in Appendix D. (Not included here)

IV *Classifications by 'tradition'.* Traditions may be defined in one or more of the following ways *inter alia.* Our purpose here is to suggest some of the bases upon which alternative perspectives rest and to draw attention to their distinctive focuses in terms of research problems and, hence in our context, distributions of documents.

(1) Bernstein[1] has identified a range of 'approach paradigms', drawing upon related work by Horton[2] and Dawe[3]:
(a) Those who place the emphasis upon the problem of order as against those who place the emphasis upon the problem of control;
(b) Those who place the emphasis upon interdependence and dependence, as against those who place the emphasis upon conflict and voluntarism;
(c) Those who place the emphasis upon how social reality is constructed out of negotiated encounters with others, and those who place the emphasis upon structural relationships;
(d) Those who emphasise the need to understand the everyday practices of members and the assumptions which make the daily practices work, and those who set up observers' categories and observers' procedures of measurement by means of which they reconstruct the constructions of members.

This points to the kinds of distinctions which are important, but requires elaboration for the practical task of organising material for literature searching. Each of these alternative emphases represents a combination of assumptions and beliefs which must be explicitly formulated as a basis for reliable classification of documents.

(2) For this purpose, Wallace[4] offers helpful ideas. His aim is 'to expose and integrate the distinctive features of several current theoretical viewpoints', bearing in mind that 'the *distinctive* features of a particular viewpoint need not be either its most *emphasized* features or its most persuasive ones'. He specifically

---

[1] Bernstein, B. Unit 17, in Swift, D.F. et al. *School and society course. E.282.* Bletchley, Open University, 1971. (Restricted circulation at time of writing.)

[2] Horton, J. 'Order and conflict theories of social problems', in *American Journal of Sociology*, 1966, 11, 701-713.

[3] Dawe, A. 'The two sociologies', in *British Journal of Sociology*, XXI(2), 207-218.

[4] Wallace, Walter L. *Sociological theory*, London, Heinemann, 1969.

excludes non-sociological thought systems. His concern is 'what kind of direct observations does each theory imply?'. He takes two foci: (a) How is the social defined? (b) How is the social explained. He stresses that 'Both questions are required, and joint answers to them will be sought'. The dual focus is crucial to avoid indefinite questions about what a given theory deals with, or what approach it represents.

His scheme is as follows:

The principal behavioral relations that define the Social are:

The principal phenomena that explain the Social are:		Objective	Subjective
Imposed on the Social via	Characteristics of the participants' environments (not-people/people)	Ecologism Demographism	(adaptive) (integrative) Functional imperativism (goal attainment) (pattern-maintenance)
	Characteristics of the participants themselves (not-nervous system/nervous system)	Materialism Psychologism	
Generated by the Social via	Characteristics of the participants' environments (not-people/people)	Technologism The social structuralisms: functional, exchange and conflict	(nonsocial objects) (social objects) Social actionism (actor) (orientation)
	Characteristics of the participants themselves (non-nervous system/nervous system)	– – Symbolic interactionism	

Definitions of the social

Objective refers to 'overt, or motor, behavior relations'.
Subjective refers to 'subjective, or covert, or dispositional behavior relations'.

Every theory is to be classified according to its prime emphasis in this respect. For operational purposes, Wallace, noting that theories emphasising subjective relations almost always involve objective relations, suggests that 'their distinguishing mark is in the way such relations are involved: such theories imply or explicitly claim that the social exists only when subjective behavior relations are inferable from, and accompany, objective relations'. This bears on, for instance, studies of social values as contrasted with social interaction. Wallace refers, as a classical example of a 'subjective behavior analysis', to Weber's definition that 'action is social insofar as, by virtue of the subjective meaning attached to it by the acting individual (or individuals), it takes account of the behavior of others and is thereby oriented in its course'.[1]

Further distinctions which may be added in are those between macro/micro levels of analysis, between genesis/maintenance of the social and between stability and change.

Explanations of the social

Imposed refers to 'conditions that are imposed on the social by the given nature of the participants or by the nature of the universe that environs them. From this point of view, the social is rather more "controlled" and "determined" by these temporally prior and/or logically more primitive conditions'. Generated refers to 'conditions that are generated by the social itself in affecting the participants or the universe that environs them. From this point of view, the social is rather more "free" or "self-determining"'.

Wallace arrived at his basic property space inductively, but notes that his dimensions are closely related to two central philosophical problems, summed up as materialism/idealism in terms of definitions, and determinism/voluntarism in terms of explanations.

Further distinctions include conditions that operate on the social through the nature of the participants themselves, which may or may not relate to the nervous system, and conditions of their environment (which may be 'people' or 'not-people' ones).

Wallace describes the nature of work in the different categories of his scheme in terms of 'isms'. So for example, he notes that the 'social structuralisms' in common with technologism are defined both in determinist terms, and as being concerned with objective behaviour relations. But 'The social structuralist viewpoints are distinguished from technologism inasmuch as the former are concerned with the things that *social participants* do to one another (no matter what material instruments they use), whereas the latter is concerned with the things that material *inventions* do to participants (no matter who originates or operates these inventions).' Both may be contrasted with 'the subjective behavior definition of the unit act in Parsons' social actionism'; also, early work of Parsons took a voluntarist position. However, the later work of Parsons and his collaborators moved away from voluntarism to functional imperativist theory, thus continuing to differ from the structuralisms on the objective/subjective dimension but moving closer to it on the imposed/generated dimension. A particular strength of the scheme is the capability it affords to show individual writers 'moving across the board' as their ideas change.

---

[1] Weber, Max Theory of social and economic organization, Glencoe, Ill., Free Press, 1947.

Wallace recognises the dysfunctions of such classification, in particular the distortion it may produce and the danger of making false distinctions. He stresses that the intention is to abstract main emphases in sociological theory only and the purpose is a practical one. In that he has considered the problems of operationalising his scheme, it offers a reasonably firm basis for classification, although the dimensions he employs are by no means the only ones upon which sociologists structure their knowledge, and the scheme fails to take account of the non-sociological work on which sociologists draw for data or ideas.

Details of the items he cites are:

Ecologism

Hawley, Amos H.. *Human Ecology*, New York, Ronald Press, 1950.

Duncan, Otis Dudley and Schnore, Leo F. 'Cultural, behavioral and ecological perspectives in the study of social organisation', in *American Journal of Sociology*, September 1959, 132-146.

Demographism

Ryder, N.B. 'Notes on the concept of a population', in *American Journal of Sociology*, March 1964, 447-463.

Materialism

Homans, George C. 'The external system', in *The human group*, New York, Harcourt, Brace and World, 1950.

Psychologism

Homans, George C. 'The institutional and sub-institutional' in Social behavior: its elementary forms, New York, Harcourt, Brace and World, 1961.

Technologism

Ogburn, William Fielding 'The hypothesis of cultural lag', in Social change with respect to culture and original nature, Ma., Peter Smith, 1964, first published 1923.

Cottrell, Fred 'Organic energy and low energy society', in Energy and society, New York, McGraw-Hill, 1955.

Functional structuralism

Merton, Robert K. 'Social structure and anomie', in American Sociological Review, 1938, 672-682.

Exchange structuralism

Blau, Peter M. 'The structure of social associations', in Exchange power in social life, New York, John Wiley, 1964.

Conflict structuralism

Van den Berghe, Pierre L 'Dialectic and functionalism: toward a theoretical synthesis', in American Sociological Review, 1963, 695-705.

Dahrendorf, Ralf 'Toward a theory of social conflict', in Conflict Resolution, 1958, 170-183.

Symbolic interactionism

Mead, George Herbert 'Play, the game, and the generalised other', in The social psychology, (edited by Anselm Strauss), Chicago, University of Chicago Press, 1956.

Blumer, Herbert 'Sociological implications of the thought of George Herbert Mead', in American Journal of Sociology, March 1966, 535-548.

Social actionism

Scott, John Finley 'The changing foundations of the Parsonian action scheme', in American Sociological Review, October 1963, 716-735.

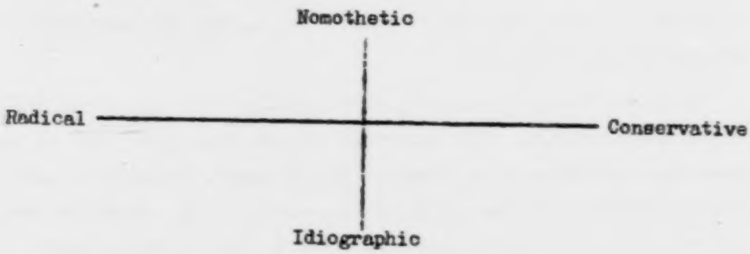
Functional imperativism

Parsons, Talcott 'Pattern variables revisited: a response of Robert Dubin', in American Sociological Review, August 1960, 467-485.

(3) A scheme of a similar kind has been outlined by Hoyle.[1] Hoyle distinguishes groups of work in the sociology of education in terms of theories concerning '"fit" between a model of society and a model of an education system'. Such theories implicitly or explicitly reflect a particular relationship between factual base, model and values. He points out that there is more to the sociology of education than the accumulation of data and that 'where the sociologist is concerned with the education-society relationship the model which he employs to bring order to the welter of data will necessarily secrete values'. This does not apply only at the macro-level. Some writers have been 'ostensibly concerned to attack "progressive" methods of education, but implicit in their criticism is a social theory which goes beyond a concern for the development of the individual child'.

Such a scheme affords meaningful grouping for material which is not strictly sociological and which, in so far as it is not excluded by the terms in which sociological frameworks are formulated, would be scattered within such schemes.

Hoyle's basic scheme is as follows:



'The terms used have the following connotations:
"Nomothetic" is used to identify those theories which focus upon the structure of education and aim to be analytical and empirical;
"Idiographic" is used to identify those theories which are concerned with cultural transmission and the content of education, and aim to be humanistic, intuitive and holistic;
"Radical" is used to identify those theories which are normatively oriented towards social and educational change of an egalitarian nature;

---

[1] Hoyle, Eric 'Social theories of education in contemporary Britain' in Social Science Information, 1970, 9(4), 169-186.

"Conservative" is used to identify those theories which are normatively oriented towards the preservation of the social and educational status quo or towards policies characterized by an ideological individualism.'

Nomothetic radicals

'The theorists who fall into this category take as their central concern the opportunity structure of society and the relationship between education and social mobility. Their basic model is technological society which is held to be characterised by change rather than stability and by only a limited integration between institutions [Floud; Halsey]. ... There exists amongst this group a certain ambivalence with regard to the functions of education in transmitting some central core of societal values or seeking to establish a common culture. ... The problem is partly a semantic one. ... It is also a substantive problem which turns upon the degree to which value transmission should be a conscious and direct aim of educational institutions or a by-product of particular forms of organization. In general the by-product view is taken here. There is an insistence that education should not be equated with socialization and the major problem is identified as the inflexibility of educational institutions.

Thus the theory centres upon the problem of wastage and has two major propositions. The empirical proposition holds that it is an economic requisite to an advanced technological society to promote the maximization of abilities at all levels. The normative proposition holds that social justice requires that no unnecessary barriers should prevent the individual from maximizing his capacities' [Official reports].

Details of the items cited are:

Halsey, A.H. 'Education and equality', in New Society, 1965, 5(142); also 'The sociology of moral education', in Niblett, W.R. (ed.) Moral education in a changing society, London, Faber, 1963.

Floud, J. 'Sociology and education', in Sociological Review Monograph 4, 1961.

15 to 18 (The Crowther report), London, HMSO, 1959.

Report of the committee under the chairmanship of Lord Robbins. (The Robbins report on higher education.) London, HMSO, 1963.

Nomothetic conservatives

'... As with the nomothetic radicals they are centrally concerned with the relationship between the opportunity structure of society and education, but their analyses and prescriptions are very different. Their basic model is that of a meritocracy and combines a theory of stratification which, if made explicit, would probably have much in common with the Davis-Moore formulation, with the belief that high intellectual capacity can and should be identified at a relatively early age and exposed to a distinctive form of education. In common with the nomothetic radicals they tend to take a "by-product" view of the socialization and cultural transmission functions of the educational system [James], although this remains very much an implicit element of this theoretical position. ... In particular the nomothetic conservatives are concerned with the preservation of the grammar school ... which ... is held to provide the freedom of choice which is not available in the state system; and in being academically selective are held to contribute to egalitarianism through providing the opportunity for poor boys, with the necessary ability to win scholarships, to enjoy the advantages of a super-elite school [Cobban]. ... A more sophisticated group of theorists who can be placed in this category are those economists whose theory of education is characterized by the advocacy of "laissez-faire" and consumer choice [West; Peacock and Wiseman].'

Details of the items cited are:

James, E. Education and leadership, London, Harrap, 1951, and other works.

Cobban, J. 'The direct grant school' in Cox, C.B. and Dyson, A.E. (eds.) Fight for education: a black paper, London, The Critical Quarterly Society, 1969.

West, E.G. Education and the state, London, Institute of Economic Affairs, 1965. 'Dr Blaug and state education: a reply', in Education: a framework for choice, London, Institute of Economic Affairs, 1967. Economics, education and the politician, London, Institute of Economic Affairs, 1968. See also Peacock, A.T. and Wiseman, J. Education for democrats, London, Institute of Economic Affairs, 1964.

Idiographic radicals

'This group of theorists approximates to the New Left political orientation. Its model of society embraces a set of institutional arrangements whereby a common culture might be achieved. Hence its educational theory is dominated by a normative theory of culture and Raymond Williams is perhaps its most outstanding exponent. ... he is an idiographic theorist since he concerns himself much more with content and tone than with the analysis of institutions. ... New Left writers would be located at a point relatively closer to the nomothetic line .... The same would probably be true of the contributors over the years to the educational journal Forum who have been largely concerned with the radical restructuring of education but have at the same time maintained an important concern with the curriculum, the nature of interpersonal relationships within the school, and the influence which a new educational style could have on the national culture. Insofar as the style is sociological it approximates more to an "educational sociology" than to a "sociology of education", the approach more characteristic of the nomothetic radicals. The same might be said of the work of the Institute of Community Studies [Jackson and Marsden] .... Perhaps the apotheosis of this approach is Michael Young's The rise of the meritocracy ....'

Details of the items cited are:

Williams, R. Culture and society, London, Penguin Books, 1961.
The long revolution, London, Chatto and Windus, 1961.
Hoare, Q. 'Education: programmes and men', in New Left Review, July-August 1965.
The work of the Institute of Community Studies, particularly the contributions of Jackson and Marsden.
Young, Michael. The rise of the meritocracy, Harmondsworth, Penguin Books, 1961.

Idiographic conservatives

'For the most part this group of theorists is composed of writers and literary critics [Eliot] .... Another key figure is the literary critic F.R.Leavis and his views have been developed within an educational context by his former students ... G.H.Bantock and William

Walsh. The model of society employed by this group is the pessimistic conception of mass society which holds that the processes of industrialization and bureaucratization have shattered social consensus and led to the alienation of the common man and the undermining of the creative powers of the elite. Their central concern is with the preservation of cultural standards, but as with the idiographic radicals there is an indiscriminate mixing of humanistic and sociological concepts of culture. In general, however, they are concerned with advocating the preservation of what are considered the highest standards in the creative arts. At the same time they are concerned with the improvement of the quality of the culture of the majority, its artefacts and its social relationships. ... The prescriptive element of the theory combines a belief in maintaining the elite-mass dichotomy with a desire to recreate a "Gemeinshaft" solidarity [see Dyson on this point] ... The educational theory of the idiographic conservatives centres upon the transmission of a particular culture. ... They reject the view that there is a widespread need for educational change .... They tend to set up for their readers a forced choice between democratization and lowered standards. ... A writer who might be included in this box, although one rather closer to the central intersection than the ones already mentioned, is Brian Wilson .... In his article on the role of the teacher, he suggests that the growing specialization of education as it enters into a closer relationship with the economy is undermining the teacher's central concern with cultural transmission. [Cf also Selznick and Clark; other work of Wilson.]

Details of the items cited are:

Eliot, T.S. Notes towards the definition of culture, London, Faber, 1948.

Leavis, F.R. 'Mass civilization and minority culture', Appendix 3 of Education and the university, London, Chatto and Windus, 1948. '"English": unrest and continuity', in Times Literary Supplement, May 29, 1969.

Bantock, G.H. Freedom and authority in education, London, Faber, 1962. Education and industrial society, London, Faber, 1963. Education and values, London, Faber, 1965. Culture, industrialisation and education, London, Routledge and Kegan Paul, 1968.

Walsh, W. The use of imagination, London, Chatto and Windus, 1955. 'Dialogue and the idea', in Cox and Dyson (eds), op. cit.

Dyson, A.E. 'The new puritanism', in Times Literary Supplement, August 19, 1960.

Wilson, B. 'The role of the teacher', in British Journal of Sociology, 1962, 13. 'The threat to university values', in Reeves, M. (ed.), Eighteen plus, London, Faber, 1966.

Selznick, P. 'Institutional vulnerability in mass society', in American Journal of Sociology, 1951, 56.

Clark, B. 'Organizational adaptation and precarious values', in American Sociological Review, 1956, 21.

Wilson, B. 'An approach to delinquency', in New Society, 1966, 7(175).

(4) The schemes both of Wallace and Hoyle identify distinctive ranges of research problems relevant to sociologists of education. Complementary to these is one which views established perspectives in relation to currently developing phenomenological perspectives, drawn from the work of Winter[1] and cited by Phillipson.[2] This identifies a new range of problems as well as presenting established areas of research in a somewhat different light.

Scientific perspectives on the social world.

Styles	Actor	Order of objects (unification)	Action (form)
Behaviourist	Impulses	Balance of forces	Laws
Functionalist	Needs	Maintenance of system	Pattern variables
Voluntarist	Interests	Domination/ compromise	Ideal constructs
Intentionalist	Attention	Order of meaning	Projects
Implicit Level	Spontaneity	Intentionality (Embodies)	Schema (body-subject)

---

[1] Winter, G. Elements for a social ethic, New York, Macmillan, 1966.

[2] Phillipson, Michael 'Theory, methodology and conceptualization', in Filmer, Paul and others, New directions in sociological theory, London, Collier-Macmillan, 1972.

(5) In the same spirit as (4), but with particular reference to emergent concerns of phenomenological sociologists working in education, the following scheme has been prepared for discussion. The aim has been to avoid the 'isms' which may mean different things to different people. We have attempted to make explicit the features of the perspectives we have characterised in their own terms, as does Hoyle. At the same time, the major dimensions come close to those of Wallace's 'explanans' and 'explanandum', although the difference in definition regroups more broadly much of the material constituting the scope of Wallace's scheme and makes a place for additional material. This does not eliminate the need for other schemes such as Wallace's, for some would reject the distinctions on which the alternative scheme is based and hence the distribution of material would not be meaningful. The latter takes account of psychological work which Wallace's does not, but does not do full justice to other non-sociological material.

NATURE OF SUBJECT MATTER

MODE OF EXPLANATION			External object-world	Members' constructions of the world
	NORMATIVE	Empirical generalisations	A	D
		Functional-means/ends	B	
	INTERPRETIVE		C	E

External object world: The social world is regarded as a reality independent of the actors who constitute it, and hence an object world in the same sense as the natural world. At the individual level this directs attention to overt behaviour and to internal functioning and states in so far as these are evidenced by overt behaviour. At the social level this means that collective phenomena are regarded as independent of the individual actors.

Constructions of the world: The social world is regarded as the externalised product of human activity, sustained or changed by this activity. As such the social world is the everyday world accomplished and experienced by its members as an organised universe of meanings in the form of typifications. Study at both individual and social level depends upon the assumption of shared understandings.

Normative: Normative explanations assume pre-determined or stable relationships (natural or teleological) amongst factors in social situations and rule governed behaviour on the part of the actor. Their validity is seen to derive from internal consistency, and their verification is by reference to data.

Interpretive: Interpretive explanations assume that social situations are self-organising and hence the notion of rule-guided (as contrasted with rule-governed) behaviour on the part of actors. This leads to concern for intersubjective validity. Verification is by reference to actors rather than to data (it should provide insights such that a stranger would be able to behave like a member and be accepted as such by other members).

For the purposes of illustration, this scheme has been operationalised in relation to a sample of journal articles for which descriptions have recently been prepared. The sample was random in that the articles were taken as they came to hand. To these have been added specific items chosen because they represent current developments in thinking which are only just beginning to be apparent in the journal literature. Descriptions of these items will be found in Appendix D. (Not included here)

A - B External object-world/Normative

A view of the social world as a world of objects external to actors, combined with a normative view of social order, defines an approach which leads to two perspectives; these are distinctive though not unrelated.

Empirical generalisations - The typical study here is that which Nagel[1] describes as seeking 'to establish relations of dependence between phenomena by correlating raw empirical data'. Phenomena thus studied may include 'non-observables' (such as mental events or what Durkheim[1] terms 'social facts'), subject to the assumption that such phenomena are to be identified externally in terms of overt behaviours. The studies in our sample include social surveys of different kinds (Holbeche; Simon; Whittaker and Watts); studies providing data on the effects of social or educational action programmes (Hunt and Hardt); correlational studies (Bezdek and Strodtbeck; Durojaiye; Entwisle and Greenberger; Scott and El-Assal); closely related are studies seeking to confirm 'Naturalist' explanations of empirically established associations of variables, eg time determination (Block, Haan and Smith; Jensen; Kinloch; Palmore, Klein and Marzuki; Ritterband; Stinchcombe, McDill and Walker), together with related critiques (Bereiter, Cronbach; Crow; Kagan; Light and Smith; Stinchcombe).

---

[1] Nagel, Ernest The structure of science, New York, Harcourt, Brace and World, 1961.

Details of the examples cited are:

Bereiter, HER, 1969, 310-318.
Bezdek and Strodtbeck, ASR, 1970, 491-502.
Block, Haan and Smith, JSI, 1969, 143-177.
Cronbach, HER, 1969, 338-347.
Crow, HER, 1969, 301-309.
Durojaiye, ER, 1969, 226-228.
Entwisle and Greenberger, SE, 1969, 238-250.
Holbeche, CE, 1969, 149-155.
Hunt and Hardt, JSI, 1969, 117-129.
Jensen, HER, 1969, 1-123.
Kagan, HER, 1969, 274-277.
Kinloch, SE, 1969, 350-367.
Light and Smith, HER, 1969, 484-510.
Palmore, Klein and Marzuki, AJS, 1970, 375-398.
Ritterband, SE, 1969, 330-349.
Scott and El-Assal, ASR, 1969, 702-709.
Simon, SR, 1969, 5-10.
Stinchcombe, HER, 1969, 511-522.
Stinchcombe, McDill and Walker, JSI, 1969, 127-136.
Whittaker and Watts, JSI, 1969, 65-89.
(For abstracts see Appendix D.) (Not included here)

---

[1] Durkheim, E. The rules of sociological method, Chicago, Chicago University Press, 1938.

Functional explanations - Nagel[1] points out that the term 'functional' is used in a variety of ways. A common element is that behavioural regularities are explained, or predicted, in terms of conformity to conditions posited to be logically necessary to ordered social existence. Conceptions of social order vary from generalised notions of 'functional prerequisites', or idealisations such as 'ideal types', to conceptions of particular concrete behaviours as goal-directed or need-fulfilling. In this sample for instance, close to Davis'[2] description of functional analysis as concerned 'to relate the parts of society to the whole, and to relate one part to another' are systems studies of various kinds (Boyle; Callahan and Robin; Cole and Adamsons; Epstein; Grinder; Kurokawa; Mumford; Warren; Weinberg and Walker). Another type of study is more practically oriented towards devising programmes of social action (Bundy; Morrison; Ritterbush; Shoben). Some work in the field of organisation theory would also be appropriately included here. We also include here studies which do not primarily deal with practical social issues but which explicitly discuss implications of their findings for social policy (Bereiter; Cronbach; Crow; Ford; Hind and Wirth; Humphreys; Jensen; Kagan; Stinchcombe, McDill and Walker).

Details of the examples cited are:

Bereiter, HER, 1969, 310-318.
Boyle, SE, 1969, 71-90.
Bundy, Daed, 1970, 531-567.
Callahan and Robin, SE, 1969, 251-260.
Cole and Adamsons, SE, 1969, 315-329.
Cronbach, HER, 1969, 338-347.
Crow, HER, 1969, 301-309.
Epstein, Daed, 1970, 700-712.
Ford, Daed, 1970, 660-675.
Grinder, JSI, 1969, 7-19.
Hind and Wirth, SE, 1969, 50-70.

---

[1] Nagel, Ernest *Op. cit.*

[2] Davis, Kingsley 'The myth of functional analysis as a special method in sociology and anthropology', in *American Sociological Review*, 1959, 24.

Humphreys, SE, 1970, 404-418.
Jensen, HER, 1969, 1-123.
Kagan, HER, 1969, 274-277.
Kurokawa, JSI, 1969, 195-213.
Morrison, Daed, 1970, 609-644.
Mumford, SR, 1970, 71-101.
Ritterbush, Daed, 1970, 645-660.
Shoben, Daed, 1970, 676-699.
Stinchcombe, McDill and Walker, JSI, 1969, 127-136.
Warren, SE, 1970, 288-310.
Weinberg and Walker, AJS, 1969, 77-96.
(For abstracts see Appendix D.) (Not included here)

C External object world/Interpretive

Studies included here are taken to regard the orderly character of the social world as different from that of the natural world (in so far as the latter implies 'an invariant sequence or concomitance of physical phenomena'[1] or 'concatenation of external forces'[2]). A functional view is also rejected in the sense of 'social order as being "out there" in an external world produced by relationships between factors external to the members of that world, primarily through the agency of shared norms and values'.[3] Social situations are seen rather as 'self organizing'[4] and the nature of social order is taken to be problematic. 'An orderly social world is established in terms of shared social meanings (or common understandings). However, whilst attention to actors' meanings distinguishes this perspective from the foregoing ones, there is a similarity of approach in that the meanings themselves are conceptualised as objects or symbols external to the actors in a situation but which the actors use in interaction with one another. Most of the studies in our sample which fall here might be said to draw upon the notion of exchange (Edgar and Warren; Green; Humphreys; Smith; Walton). Another type of study is that concerned with the 'deep structure' of relationships (Young). Much work drawing upon ideas of symbolic interaction would be appropriately placed in this context.

---

[1] Walsh, David 'Sociology and the social world', in Filmer, Paul and others, Op. cit.

[2] Walsh, David 'Varieties of positivism', in Filmer, Paul and others, Op. cit.

[3] Walsh, David 'Sociology and the social world' in Filmer, Paul and others, Op. cit.

[4] Filmer, Paul 'On Harold Garfinkel's ethnomethodology', in Filmer, Paul and others, Op. cit.

Details of the examples cited are:

Edgar and Warren, SE, 1969, 386-399.
Green, HER, 1969, 221-252.
Humphreys, SE, 1970, 404-418.
Smith, ASR, 1970, 860-873.
Walton, AJS, 1970, 828-851.
Young, ASR, 1970, 297-307.
(For abstracts see Appendix D.) (Not included here)

D Constructions of the world/Normative

Studies here characterise social situations in terms of 'the general principles according to which man in daily life organises his experiences' (Schutz[1]). A contrast is drawn between the social world and the natural world. Social meanings cannot be taken as 'givens'. Meaning is not then a fixed property of objects or situations. Social meanings reflect actors' assumptions about the phenomena comprising the social world. Such assumptions are problematic and become the focus of sociological investigation. The studies here are taken to characterise (by contrast with the studies in the following section which tend to centre on the processes of) the social construction of reality. The problematic focus varies across actors' differing perceptions and knowledge of the social world (Clark & Clark; Friedenberg; Metzger; Sandford); observer/actor differences in perspective (Slater); implicit perspectives introduced by the observer (Barrett; Deutsch; Thompson; Troutner). Similar notions underlie studies of labelling (Warren), together with some work in other substantive areas not represented in our sample, such as stereo-typing. Some work in the Marxist tradition and in ethnomethodology would also have a place here.

---

[1] Schutz, A. Collected papers, volume 1, The Hague, Nijhoff, 1962.

Details of examples cited are:

Barrett, SR, 1969, 251-266.
Clark and Clark, SE, 1969, 293-314.
Deutsch, JSI, 1969, 5-18.
Friedenberg, JSI, 1969, 21-38.
Metzger, Daed, 1970, 568-608.
Sandford, JSI, 1969, 137-146.
Slater, ASR, 1969, 359-373.
Thompson, BJES, 1970, 18-31.
Troutner, HER, 1969, 124-154.
Warren, NS, 1969, 321-322.
(For abstracts see Appendix D.) (Not included here)

E Constructions of the world/Interpretive

Here again we find the assumption of a qualitative difference between the natural and the social orders, and the focus upon the social construction of reality. It is assumed also that this socially constructed reality is an ongoing accomplishment: '... rules are enacted in social situations through a continuous interpretation of their meanings in the context of commonsense decision-making'.[1] We lack, in our sample, general studies of the processes involves. Such studies would be expected, as do Berger and Luckman,[2] to assume a dialectic process between two types of reality: objective reality refers to typifications which are the outcome of people interacting with one another in routine ways and which come to be regarded as objectively real; subjective reality concerns individual experiences as people are confronted by and apprehend this so-called objective reality. Studies in our sample include discussion of what this perspective implies for the sociological analysis of education (Davies), as well as application to educational situations (Bernstein, Esland, Keddie). It is perhaps worth special mention that the construction of sociological explanations, as well as explanations constructed by actors, is regarded as a problematic feature in the investigation of social situations (Blum).

---

[1] Walsh, David. Op. cit.

[2] Berger, P.L. and Luckmann, T. The social construction of reality: a treatise in the sociology of knowledge, Harmondsworth, Penguin Books, 1967. (First published in U.S.A. in 1966.)

Details of items cited are:

Bernstein, Young (ed.), 1971.
Blum, Young (ed.), 1971.
Davies, Soc, 1970, 1-22.
Esland, Young (ed.), 1971.
Keddie, Young (ed.), 1971.
(For abstracts see Appendix D.) **(Not included here)**

(6) On rather different lines would be classification in terms of 'significant others', or in other words traditions defined in terms of people. Such a scheme has certain drawbacks. In the first place, there is a variety of reasons for which a researcher may cite the work of others, some perhaps of somewhat dubious validity. Problems, arising from this variety, may be overcome by working from the text and responding only to those others on whose work the researcher has actually built, rather than to all those he cites.[1] In the second place, the utility of such a scheme is dependent on the user's familiarity with the literature so as to select an appropriate author as a starting point. Its advantage lies in identifying what may be called mini-traditions (eg work centring round Coleman's investigation of the school as a social system by means of his subcultural model), and affording a grouping of material which may aid consideration of appropriateness of alternative perspectives on given substantive problems.

---

[1] Lipetz, B. 'Improvement of the selectivity of citation indexes to science literature through inclusion of citation relationship indicators', in American Documentation, 1965, 16(2), 81-90.

The following schema indicates how this might be handled:

(a) Theories or models developed by 'other' and employed in work being classified.	Arrangement by name of 'other' and date of source work.
(b) Specific empirical questions raised in work of 'other' and investigated in work being classified.	
(c) Specific methods developed by 'other' and employed in work being classified.	

(7) Concluding this section on a more general note, it will be evident that the substructure of perspectives in all these schemes is often dictated to a large extent by the perspectives themselves. For instance there is clearly a broad difference between 'object-world' and 'construction' perspectives in that the former appear to lend themselves to taxonomising and, by virtue of relative stability, allow of more detailed sets of categories; the latter on the other hand necessarily concern views which are more fluid and looser knit. In all cases, as new research problems are opened up, so new subsets of material may be formed. Similarly new perspectives may be distinguished as they emerge and established ones may dwindle away or become reoriented. Reworking of the material will be necessary from time to time to take account of such developments. It seems improbable that detailed subdivision or elaborate cross-referencing amongst perspectives will be appropriate. Users' conceptual strategies will guide us on this point. In general it is expected that mutual definition of subsets of material within a category, and mutual definition of perspectives one of another, will suggest alternative directions of search without prescribing them.

V Classification in terms of operational aspects of research. This series of classifications is seen to relate to a mode of search in which the researcher starts from a specific research question and seeks a perspective which will offer helpful insights and provide a language in which to think about his problem. These classifications are intended to accord with the stage of research at which the researcher abstracts the salient features of the situation under investigation in order to translate his problem into operational terms. Not all the material included in the system is relevant to this kind of work.

We draw here upon Riley[1], who suggests a multi-dimensional scheme for classifying research designs in terms of the 'series of basic choices the researcher makes in planning the types of data to be assembled and organised and the types of procedures to be used in his study'. Documents may be classified in terms of as many options as apply. Researchers may classify their problems in a similar way. The kinds of options, or the kinds of questions to be asked, are seen to represent factors taken into account in the literature searching process. Different search strategies would be expected to emphasise some options at the expense of others and to reflect different orders of priority amongst options.

Riley conceives of these options in broad terms. For our purpose more detail is required. Nevertheless, the categories we establish must remain relatively broad. The interchangeability of operational indicators[2] is such that we have to think at the level of the properties to be represented in the research design rather than the indicators used to measure or characterise them. Only a data bank can serve a researcher who wishes to start from data and build up; a system which takes the document as the basic unit has to gear itself to the model which guides operations such as data collection.

From a classification of his problem in terms of the factors which guide decisions concerning operational investigation, the system should allow the researcher to proceed in one of two directions. He may be directed to one or more of the perspectives represented in the earlier set of classifications. Alternatively he may prefer to go directly to the abstracts in the total listing of items in the system and make an independent appraisal of them in terms of perspective and approach relative to his problem.

Riley's scheme refers to options presented by the following considerations:

I Nature of research case (level of analysis).

II Number of cases.

---

[1] Riley, Matilda White Sociological research. I A case approach, New York, Harcourt, Brace and World, 1963.

[2] For example length of education may be taken as an indicator of either social class or achievement, as well as being of possible interest for its own sake.

III Sociotemporal context.

IV Primary basis for selecting cases (sample as representational/analytic of conceptual universe).

V Time factor (static/dynamic).

VI Extent of researcher's control over the system under study (ranging from field study to laboratory study).

VII Basic sources of data (already available/collected for purpose in hand).

VIII Method of gathering data (observation/questioning).

IX Number of properties used in research.

X Method of handling single properties (quantitative/qualitative).

XI Method of handling relationships amongst properties (quantitative/qualitative).

XII Treatment of properties (unitary/collective).

Taken together, these dimensions all represent a particular objectivist perspective. We have selected, modified and elaborated within this generalised framework in relation to the literature searching process. Decisions upon some options are hardly likely to prompt recourse to the literature. Some require careful specification for the classification of documents in that certain points tend to be left unclear in the reporting of research (eg the population which a given sample is taken to represent, and hence the generalisability of the study). Greater detail is required in that, whilst not all options are relevant to all documents, a simple dichotomisation is not likely to save the searcher much time. Additionally, there must be freedom to take account of conceptualisation which has guided the author's work. It should be noted that the classification of a document is not determined by the use of particular words in an author's account of his work but by the meaning the author attaches to the words he uses. For example, two authors may speak of student protest but the concern of one might be with a distinctive pattern of values and of the other with actual protest behaviours.

At this stage, we consider alternative dimensions separately, with the object of promoting discussion amongst subject experts as to their collective and individual validity and their utility in the literature searching context. The further detail to be built in, and the ways in which these dimensions might be interlinked, or linked with those in the earlier set of classifications, pends on investigation of the nature of the search strategies to be supported by the system.

Dimensions I, III, IV, VIII, IX, X, XI of Riley's scheme have provided ideas particularly relevant to our purpose, and suggest the following set of classifications. (Examples are drawn from the sample of items used to illustrate scheme 5 in the earlier set of classifications (see section IV of this appendix) to illustrate how different categories are defined. A few items have been excluded as inappropriate for treatment in such terms.)

A Case for study/level of analysis

Individual

Relationship between individual and social

Social

These categories may be defined as follows:

Individual

Defined as including relationships amongst, or dimensions of, individual characteristics (cf Jensen, HER, 1969, 1-123), together with the individual's experiences of the social (cf Block, Haan and Smith, JSI, 1969, 143-177). This category is extended to include studies of particular roles, and studies of status sets associated with particular social positions (cf Cole and Adamsons, SE, 1969, 315-329).

Relationship between individual and social

Emphasis is upon the individual in interaction with the social (cf Mumford, SR, 1970, 71-101), as contrasted with 'individual' which corresponds rather to Lukes'[1] characterisation of work in which 'the relevant features of the social context are, so to speak, built into the individual'. The boundary of the category 'relationship between individual and social' is drawn so as to include studies of individual/individual as well as individual/group relationships.

---

[1] Lukes, Steven 'Methodological individualism reconsidered', in British Journal of Sociology, 1968, XIX(2), 119-129.

Social

This category is defined to include interaction within (cf Warren, SF, 1970, 288-310) and across groups (cf Weinberg and Walker, ASJ, 1969, 77-96). We make no attempt to distinguish between (a) a given pattern of interaction or unit studied as a subset of a more inclusive set and (b) the same interaction or unit studied as itself a set inclusive of subsets.

The boundaries between such categories could be drawn differently, and may require adjustment in practice. Finer distinctions are often made in terms of system levels. However, in a practical situation, the attempt to make such further distinctions is likely to have low reliability, and we do not make it.

This dimension underlies later ones and we do not therefore illustrate it here.

B Universe of study

This refers to what Riley terms 'the conceptual universe' of study. Whilst analytically distinct, it is not very meaningful to consider this dimension in isolation from the previous (case for study) one. Sampling is guided by the conceptual universe to be studied. Characterisation of the conceptual universe is thus interdependent with level of analysis. The options may therefore be presented as follows:

Individual, defined in terms of
- Ascribed characteristics;
- Achieved characteristics;
- Formal position in social structure.

Individual in relation to social
- as above, and/or defined by the social, depending on focus of study.

Social, defined in terms of
- Basic social units (eg peer group, formal organisation);
- Basic social institutions (eg education, economy, polity as institutionalised agencies for fundamental social processes such as socialisation, social control).

By employing dimensions A and B, we get the following distribution of material in our sample (items are sorted into as many types of category as apply):

Case for study/Universe of study

* Individual in relation to social indicated thus.

All categories broadly defined, detailed subdivision as required.

Individual

Ascribed characteristics

Ethnic or racial (Bereiter, HER, 1969, 310-318; Cole and Adamsons, SE, 1969, 315-329; Cronbach, HER, 1969, 338-347; Crow, HER, 1969, 301-309; Durojaiye, ER, 1969, 226-228; Entwisle and Greenberger, SE, 1969, 238-250; Hunt and Hardt*, JSI, 1969, 117-129; Jensen, HER, 1969, 1-123; Kagan, HER, 1969, 274-277; Kurokawa*, JSI, 1969, 195-213; Light and Smith, HER, 1969, 484-510; Ritterband*, SE, 1969, 330-349; Stinchcombe, HER, 1969, 511-522); Sex (Bezdek and Strodtbeck, ASR, 1970, 491-502).

Achieved characteristics

Educational affiliation: ability category (Hind and Wirth*, SE, 1969, 50-70; Kinloch*, SE, 1969, 350-367); Occupational affiliation: educational groups including academics (Cole and Adamsons, SE, 1969, 315-329; Deutsch, JSI, 1969, 5-18; Edgar and Warren*, SE, 1969, 386-399; Humphreys*, SE, 1970, 404-418; Simon, SR, 1969, 5-10); Occupational affiliation: other groups (Kinloch*, SE, 1969, 350-367; Mumford*, SR, 1970, 71-101); Political affiliation (Cole and Adamsons, SE, 1969, 315-329); Socio-economic affiliation (Bereiter, HER, 1969, 310-318; Cole and Adamsons, SE, 1969, 315-329; Cronbach, HER, 1969, 338-347; Crow, HER, 1969, 301-309; Entwisle and Greenberger, SE, 1969, 238-250; Jensen, HER, 1969, 1-123; Kagan, HER, 1969, 274-277; Kinloch*, SE, 1969, 350-367; Light and Smith, HER, 1969, 484-510; Ritterband*, SE, 1969, 330-349; Stinchcombe, HER, 1969, 511-522); Non-affiliation (Whittaker and Watts, JSI, 1969, 65-89).

Formal characteristics

Role differentiation (Bezdek and Strodtbeck, ASR, 1970, 491-502; Callahan and Robin, SE, 1969, 251-260); Reciprocal roles: power relations (Edgar and Warren*, SE, 1969, 386-399; Smith*, ASR, 1970, 860-873); Status sets (Cole and Adamsons, SE, 1969, 315-329).

Social

Social units

Dyad (Smith*, ASR, 1970, 860-873); Group: Deviant group (Warren, NS, 1969, 321-322; Whittaker and Watts, JSI, 1969, 65-89), Ethno-religious group (Barrett, SR, 1969, 251-266; Kurokawa, JSI, 1969, 195-213; Palmore, Klein and Marzuki, AJS, 1970, 375-398; Slater, ASR, 1969, 359-373; Stinchcombe, McDill and Walker, JSI, 1969, 127-136), Generation (Friedenberg, JSI, 1969, 21-38; Grinder, JSI, 1969, 7-19; Sandford, JSI, 1969, 137-146), Peer group (Grinder, JSI, 1969, 7-19; Warren, SE, 1970, 288-310), Political group (Block, Haan and Smith, JSI, 1969, 143-177; Clark and Clark, SE, 1969, 293-314; Ford, Daed, 1970, 660-675; Metzger, Daed, 1970, 568-608; Scott and El-Assal, ASR, 1969, 702-709; Weinberg and Walker, AJS, 1969, 77-96), Reference group (Clark and Clark, SE, 1969, 293-314; Edgar and Warren*, SE, 1969, 386-399; Sandford, JSI, 1969, 137-146), Socio-economic group (Slater, ASR, 1969, 359-373); Organisation (Edgar and Warren*, SE, 1969, 386-399; Mumford*, SR, 1970, 71-101).

We believe there may be further dimensions to be considered in the classification of social units. For instance, Young (ASR, 1970, 297-307) and related items appear to represent a conception of social unit unlike that which is implicit in the headings above.

Social institutions

Economy and Family & Kinship (Palmore, Klein and Marzuki, AJS, 1970, 375-398); Economy and Polity (Barrett, SR, 1969, 251-266; Walton, AJS, 1970, 828-851); Education (Doyle, SE, 1969, 71-90; Bundy, Daed, 1970, 531-567; Clark and Clark, SE, 1969, 293-314; Edgar and Warren*, SE, 1969, 386-399; Ford, Daed, 1970, 660-675; Hind and Wirth*, SE, 1969, 50-70; Humphreys*, SE, 1970, 404-418;

Hunt and Hardt*, JSI, 1969, 117-129; Metzger, Daed, 1970, 1970, 568-608; Morrison, Daed, 1970, 609-644; Ritterband*, SE, 1969, 330-349; Scott and El-Assal, ASR, 1969, 702-709; Stinchcombe, McDill and Walker, JSI, 1969, 127-136; Warren, SE, 1970, 288-310); *Education and Economy* (Kinloch*, SE, 1969, 350-367; Ritterbush, Daed, 1970, 645-660; Shoben, Daed, 1970, 676-699); *Education and Polity* (Epstein, Daed, 1970, 700-712; Green, HER, 1969, 221-252; Weinberg and Walker, AJS, 1969, 77-96).

C Sociotemporal context

The need for a *temporal* dimension does not occur with our present sample, but could readily be built in to the system. *Geographical setting* is another dimension which is relevant under this heading. This could be organised as a regional presentation of countries, or could be tied to more general distinctions such as advanced/developing, Western/non-Western. Considerations which are important to those engaged in cross-cultural analysis would decide the matter of presentation. Sample details will provide the relevant data.

The geographical spread represented by our current sample is as follows:

Canada (Humphreys, SE, 1970, 404-418; Kurokawa, JSI, 1969, 195-213); *France* (Clark and Clark, SE, 1969, 293-314; Weinberg and Walker, AJS, 1969, 77-96); *Latin America* (Walton, AJS, 1970, 828-851; Weinberg and Walker, AJS, 1969, 77-96); *Malaysia* (Palmore, Klein and Marzuki, AJS, 1970, 375-398); *Nigeria* (Barrett, SR, 1969, 251-266); *UK* (Durojaiye, ER, 1969, 226-228; Holbeche, CE, 1969, 149-155; Weinberg and Walker, AJS, 1969, 77-96); *US* (Block, Haan and Smith, JSI, 1969, 143-177; Doyle, SE, 1969, 71-90; Bundy, Daed, 1970, 531-567; Callahan and Robin, SE, 1969, 251-260; Cole and Adamsons, SE, 1969, 315-329; Edgar and Warren, SE, 1969, 386-399; Entwisle and Greenberger, SE, 1969, 238-250; Epstein, Daed, 1970, 700-712; Grinder, JSI, 1969, 7-19; Hind and Wirth, SE, 1969, 50-70; Hunt and Hardt, JSI, 1969, 117-129; Kinloch, SE, 1969, 350-367; Metzger, Daed, 1970, 568-608; Ritterband, SE, 1969, 330-349; Scott and El-Assal, ASR, 1969, 702-709; Shoben, Daed, 1970, 676-699; Slater, ASR, 1969, 359-373; Smith, ASR, 1970, 860-873; Stinchcombe, McDill and Walker, JSI, 1969, 127-136; Warren, SE, 1970, 288-310; Weinberg and Walker, AJS, 1969, 77-96; Whittaker and Watts, JSI, 1969, 65-89).

Another aspect of sociotemporal context is that of social context, as contrasted with geographical context. Subject experts have impressed upon us the importance of the distinction between everyday and technical terminology in the definition of research problems. There may be a need to provide a set of entry points for a stage of search prior to that at which the problem for investigation has been formulated in sociological terms. We may again draw upon details of sample to specify the individuals or social units actually studied, as contrasted with the conceptual universe they are chosen to represent.

A particular problem with our material is the use of a given word with both a 'lay' and a 'technical' meaning. Ideally one would specify the conceptual universe to which the findings of a study are validly generalisable in both lay and technical terms to cater for different modes of access. In practice, this is an impossible task. It seems a reasonable second best for the 'lay' approach to equate a social (as contrasted with a sociological) definition of problem with the social setting. This may be the research subjects or the physical setting in which a study was carried out. For this purpose we may again draw upon sample data, 'rounding up' to broad categories (eg individuals to educational/developmental categories). Relatively arbitrary 'rounding up' seems inevitable because it is frequently unclear whether a sample of individuals is selected for study by virtue of, say, being secondary school pupils, by virtue of race or as being disadvantaged because of race. Again, a school may be the setting of a study not for interest in its own right but as a 'convenient' setting in which to study the functioning of a formal organisation.

Such a classification would be an ancillary one, and would attempt no distinction between the convenience sample and the sample which closely reflects the conceptual universe of a study. The value of this type of classification may be assessed from the following illustration:

Research subjects

Defined by combined educational/development characteristics

Infants and children (at pre-education stage)

—

Children (at primary education stage)

Durojaiye, ER , 1969, 226-228; Entwisle and Greenberger, SE, 1969, 238-250.

Young people (at secondary education stage)

Bozdek and Strodtbeck, ASR, 1970, 491-502; Boyle, SE, 1969, 71-90; Callahan and Robin, SE, 1969, 251-260; Grinder, JSI, 1969, 7-19; Kurokawa, JSI, 1969, 195-213; Smith, ASR, 1970, 860-873.

Young people and adults (in post secondary education)

Block, Haan and Smith, JSI, 1969, 143-177; Hind and Wirth, SE, 1969, 50-70; Holbeche, CE, 1969, 149-155; Hunt and Hardt, JSI, 1969, 117-129; Ritterband, SE, 1969, 330-349; Scott and El-Assal, ASR, 1969, 702-709; Smith, ASR, 1970, 860-873; Weinberg and Walker, AJS, 1969, 77-96; Whittaker and Watts, JSI, 1969, 65-89.

Young people and adults (in role other than that of educand)

Education as occupation

Administrators: Edgar and Warren, SE, 1969, 386-399; Teachers - general: Edgar and Warren, SE, 1969, 386-399; primary/secondary: Humphreys, SE, 1970, 404-418, Warren, SE, 1970, 288-310; post-secondary including research function: Cole and Adamsons, SE, 1969, 315-329, Simon, SR, 1969, 5-10.

Other roles and occupations

Administrators: Walton, AJS, 1970, 828-851; Dropout: Whittaker and Watts, JSI, 1969, 65-89; Parent and familial: Kurokawa, JSI, 1969, 195-213, Palmore, Klein and Marzuki, AJS, 1970, 375-398; Peace Corps: Block, Haan and Smith, JSI, 1969, 143-177; Professionals (other than teachers) Kinloch, SE, 1969, 350-367, Mumford, SR, 1970, 71-101.

Research setting (defined as formal setting of study)

Educational framework

Preschool: —; Primary: Stinchcombe, McDill and Walker, JSI, 1969, 127-136; Secondary: Boyle, SE, 1969, 71-90, Callahan and Robin, SE, 1969, 251-260, Stinchcombe, McDill and Walker, JSI, 1969, 127-136; Post secondary: Hind and Wirth, SE, 1969, 50-70, Holbeche, CE, 1969, 149-155, Scott and El-Assal, ASR, 1969, 702-709; Special programmes - compensatory: Hunt and Hardt, JSI, 1969, 117-129.

Other

Industrial firms: Kinloch, SE, 1969, 350-367, Mumford, SR, 1970, 71-101; Research centres: Warren, NS, 1969, 321-322.

D Properties

An alternative way of characterising the content of a study is in terms of the properties selected for investigation. Riley terms this 'the representation in research of social system properties'. The case for study is here defined in terms of the sum of its salient properties.

We identify the following categories:

Orientations

eg attitudes, beliefs, frames of reference, goals, values;

Performances

eg qualities such as achievement, effectiveness, intelligence, productivity, together with concrete behaviours such as collective behaviour;

Structure and functioning

eg acculturation, assimilation, bureaucracy, complexity, differentiation, identification, socialisation, stratification.

Each would include both work where the property is studied in relation to others and also work aimed at delineating the property in terms of its dimensions, eg pragmatic/idealistic value orientations.

There is a practical overlap amongst these categories even whilst they are analytically distinct. Orientation is sometimes taken as an indicator of behaviour, and inferences about orientations may be made from reported or observed behaviours. Behaviours may be described in terms of underlying social structure or process, and social structure or process may be characterised by reference to concrete behaviours. There is, too, considerable conceptual overlap within categories because of interchangeability of indicators. Thus no valid distinction can be made between, say, achievement and intelligence as these terms are employed in reporting upon the collection and interpretation of data in the literature.

These do not form mutually exclusive types of categories into which a total study may be classified as with the research case. One is here classifying separate elements of a study or variables. It is possible to argue that it is not meaningful to think in terms of variables divorced from context. It seems again necessary to take account of level of analysis at least. We also feel some reservation about undertaking this task in isolation from perspective*, even with the contents of the categories enumerated only at a very general level. A given term such as social class may, as used, refer to either orientation, performance or structure and functioning. Compare, for instance, social class as it features in Slater (ASR, 1969, 359-373) and as it features in Jensen (HER, 1969, 1-123). Other ways of categorising variables which have been considered but are not illustrated because they were found more difficult to make than that suggested here (eg structure/process) present a similar problem. It seems that implicitly at least, in considering a named variable to decide what the word 'means' as a variable, one must take account of the author's perspective. It may be that the different conceptual bases of these categories should be made explicit.

There is also the practical point that the searcher may be interested in a combination of variables, ie in the total study rather than in a single element. Here we assume that his dependent variable may be taken as a starting point. Thus, by distinguishing the status (eg dependent/independent, higher order concept, main focus) of a given variable in a given study, we may reduce the number of items to be scanned for a particular combination of variables. With a large number of items in the system, it would be possible as standard practice to subdivide material associated with any variable by 'other variable(s)'. We return to this point later.

Taking account of perspective only in so far as perspective is associated with level of analysis, and sorting material in terms of categories of variables in the way described, our sample produces the following distribution of material. We present this for consideration in the light of the problems which we have suggested are inherent in such a form of analysis:

---

* Considerations of perspective are clearly not absent from the other classifications in this set, but the risk of scrambling perspectives seems to be much greater when studies are analysed into variables and these variables (rather than total studies alike in some respect) are grouped into categories; such categories represent a range of decontextualised meanings.

Properties

* Individual in relation to social indicated thus.

All categories broadly defined, detailed subdivision as required.

Orientations (in terms of object of orientation)

Individual

General (Bozdok and Strodtbeck, ASR, 1970, 491-502; Hunt and Hardt*, JSI, 1969, 117-129); Education (Holbeche, CE, 1969, 149-155; Humphreys*, SE, 1970, 404-418; Ritterband*, SE, 1969, 330-349); Occupation (Hind and Wirth*, SE, 1969, 50-70); Politics (Cole and Adamsons, SE, 1969, 315-329); Race (Durojaiye, ER, 1969, 226-228).

Social

General (Kurokawa*, JSI, 1969, 195-213); Education (Bundy, Daed, 1970, 531-567; Clark and Clark, SE, 1969, 293-314; Ford, Daed, 1970, 660-675; Friedenberg, JSI, 1969, 21-38; Green, HER, 1969, 221-252; Grinder, JSI, 1969, 7-19; Metzger, Daed, 1970, 568-608; Morrison, Daed, 1970, 609-644; Sandford, JSI, 1969, 137-146; Slater, ASR, 1969, 359-373); Occupation (Mumford*, SR, 1970, 71-101); Politics (Block, Haan and Smith, JSI, 1969, 143-177); Self (Slater, ASR, 1969, 359-373); Society (Deutsch, JSI, 1969, 5-18); Individual differences (Stinchcombe, McDill and Walker, JSI, 1969, 127-136; Warren, NS, 1969, 321-322).

Performances

Individual

Behaviours: Educational (Callahan and Robin, SE, 1969, 251-260; Ritterband*, SE, 1969, 330-349); Leadership choice (Callahan and Robin, SE, 1969, 251-260; Durojaiye, ER, 1969, 226-228); Political (Cole and Adamsons, SE, 1969, 315-329); Research (Simon, SR, 1969, 5-10).

Qualities: Achievement, intelligence etc (Bereiter, HER, 1969, 310-318; Cronbach, HER, 1969, 338-347; Crow, HER, 1969, 301-309; Durojaiye, ER, 1969, 226-228; Hunt and Hardt*, JSI, 1969, 117-129; Jensen, HER, 1969, 1-123; Kagan, HER, 1969, 274-277; Kinloch*, SE, 1969, 350-367; Light and Smith, HER, 1969, 484-510; Stinchcombe, HER, 1969, 511-522).

Careers: Mobility (Holbeche, CE, 1969, 149-155; Kinloch*, SE, 1969, 350-367).

Social

Behaviours: Collective (Block, Haan and Smith, JSI, 1969, 143-177; Scott and El-Assal, ASR, 1969, 702-709; Weinberg and Walker, AJS, 1969, 77-96; Young, ASR, 1970, 297-307); Decision making (Walton, AJS, 1970, 828-851); Educational (Ritterbush, Daed, 1970, 645-660; Shoben, Daed, 1970, 676-699).

Careers: Mobility (Slater, ASR, 1969, 359-373).

Structure and functioning

Individual

Biological: Genetic characteristics (Bereiter, HER, 1969, 310-318; Cronbach, HER, 1969, 338-347; Crow, HER, 1969, 301-309; Jensen, HER, 1969, 1-123; Kagan, HER, 1969, 274-277; Light and Smith, HER, 1969, 484-510; Stinchcombe, HER, 1969, 511-522).

Psychological including linguistic: Experience (Block, Haan and Smith, JSI, 1969, 143-177; Whittaker and Watts, JSI, 1969, 65-89); Language (Entwisle and Greenberger, SE, 1969, 238-250); Personality (Durojaiye, ER, 1969, 226-228; Edgar and Warren*, SE, 1969, 386-399; Kurokawa*, JSI, 1969, 195-213; Mumford*, SR, 1970, 71-101; Smith*, ASR, 1970, 860-873; Whittaker and Watts, JSI, 1969, 65-89); Sex role (Bezdek and Strodtbeck, ASR, 1970, 491-502).

Sociological: Role and status (Callahan and Robin, SE, 1969, 251-260; Cole and Adamsons, SE, 1969, 315-329; Edgar and Warren*, SE, 1969, 386-399; Humphreys*, SE, 1970, 404-418; Smith*, ASR, 1970, 860-873).

Social (in terms of social phenomena defined as both process and effect of process)[≠]

Social change (including planned change) (Barrett, SR, 1969, 251-266; Deutsch, JSI, 1969, 5-18; Epstein, Daed, 1970, 700-712; Green, HER, 1969, 221-252; Hunt and Hardt*, JSI, 1969, 117-129; Palmore, Klein and Marzuki, AJS, 1970, 375-398; Ritterbush, Daed, 1970, 645-660; Shoben, Daed, 1970, 676-699; Walton, AJS, 1970, 828-851); Social control (Warren, SE, 1970, 288-310);

[≠] It is unclear whether the papers in Young's Knowledge and control symposium would be appropriately included in this section.

Social cont'd

Socialisation (Edgar and Warren*, SE, 1969, 386-399; Smith*, ASR, 1970, 860-873; Warren, SE, 1970, 288-310); Social integration (Kurokawa*, JSI, 1969, 195-213; Mumford*, SR, 1970, 71-101; Young, ASR, 1970, 297-307); Stratification (Barrett, SR, 1969, 251-266; Boyle, SE, 1969, 71-90; Hind and Wirth*, SE, 1969, 50-70; Humphreys*, SE, 1970, 404-418; Kinloch*, SE, 1969, 350-367; Ritterband*, SE, 1969, 330-349; Scott and El-Assal, ASR, 1969, 702-709; Stinchcombe, McDill and Walker, JSI, 1969, 127-136; Warren, NS, 1969, 321-322).

Experts have suggested that there is a place for an additional schema which focusses on variables at the level of the total study rather than handling the variables in a given study independently of each other. Such a schema might refer to either combinations of variables or the research area to which a study makes a contribution. In the first case, each variable would be presented in combination with all the other variables it is studied with. In the second case, the schema would be restricted to combinations of variables, controversies (eg heredity/environment) etc, which represent established research areas and hence might be sought by the searcher.

Both are possible, but neither is likely in itself to be practically helpful. A particular complicating factor with the first is the interchangeability of indicators mentioned earlier, and the number of variables with which a given variable is intertwined. A further factor is that the same problem may be stated at a number of levels of generality, and a combination of levels may be an essential feature of a given study. It is also possible that the same word may be used at different levels. There is a related problem when we think in terms of research areas. Such areas tend to focus on a particular range of research questions but are not mutually exclusive in terms of the phenomena studied. For example, study of almost any substantive area would provide an example of innovation; almost any sociological theory would yield some insights into a given educational problem.

For these kinds of reasons, with either type of scheme, it is impossible to define headings which designate areas which are distinctive and mutually exclusive, and the amount of overlap of different kinds is likely to be such that the amount and complexity of cross-referencing required would defeat the patience of the user, even if it were possible to clarify the links to be made. It seems that a scheme intended to handle the total study must take explicit account of 'slant', in the sense of the kinds of questions to be answered, as well as of the phenomena to which the questions are addressed. For instance, in our present sample we find items in which findings are intended variously to validate models (Callahan and Robin, SE, 1969, 251-260; Kinloch, SE, 1969, 350-367), to provide insights into the functioning of social institutions (Mumford, SR, 1970, 71-101), to guide policy making (Boyle, SE, 1969, 71-90). We would welcome further suggestions as to the range of considerations which should guide the building of a framework within which classification in terms either of combinations of variables or research areas will be meaningful.

E Author and author's affiliation

This does not seem to require discussion.

Conclusion

It seems clear that only rarely will a single one of these dimensions express all the aspects of a problem under investigation. We have stressed that, as with the previous set of classifications, it is our purpose at this stage merely to disentangle the various dimensions taken into account prior to the operational stage of research and to select those relevant in literature searching. No claim is made yet to have identified all the relevant ones. For instance, it is possible that for some searches the most direct approach to a particular kind of research question may be one in terms of research method (sociometry, factor analysis), or even instrument (semantic differential). It will be possible to marry dimensions as in the previous set of schemes; generalised types of search strategies will determine how they may most usefully be combined. There is clearly a limit to the amount of complexity which can be built into a printed page index. Ultimately, computer manipulation of the index would allow for much greater variety and complexity in the ways in which a searcher could plan and modify a search as it progressed.

Coversheet to:
APPENDIX D to
Paper III 6B (iii)

Samples of items referred to in sections IV and V of Appendix C: Bibliographical details and description of contents

ABBREVIATIONS

AJS	American Journal of Sociology
ASR	American Sociological Review
BJES	British Journal of Educational Studies
CE	Comparative Education
Daed.	Daedalus
ER	Educational Research
HER	Harvard Educational Review
JSI	Journal of Social Issues
NS	New Society
SR	The Sociological Review
Soc.	Sociology
SE	Sociology of Education

III 6B. (iv) Procedures in constructing a scheme for the intellectual organisation of documents in terms of the dimensions by which knowledge comes to be structured, December, 1972

An account of this aspect of our work necessarily takes the form of an interim report, together with ideas for future investigation. Our purpose is to construct a scheme in which the notion of perspective is employed as the organising concept. The function of the scheme is to allow literature searching to proceed in a way which allows of knowledge generation, and effectiveness is defined in terms of non-interference with this process.

1. Preliminary steps

Consultation with subject experts made it clear that the notion of perspective is one which is not only meaningful but also necessary to subject experts in our field. It was argued that, so long as the notion is central to sociological discussion, it will be central in the knowledge generation process. Effective realisation of the notion in a literature searching system, however, is not simply a question of reflecting in the system the full complexity of perspectives employed in the field. It depends on two main factors: (a) operationalising the concept for our purpose, ie in such a way that it can be employed systematically rather than intuitively (and hence often differently) by both system operators and system users; (b) operationalising it in a way which does minimum violence to it, ie in relation to actual perspectives and how they shape sociological thinking. We have to find a compromise between these technical and conceptual considerations.

The most recent phase of our work has been addressed to both these points. A survey of the theoretical literature led to a preliminary outline and illustration of a method for characterising perspectives in the context of a literature searching system, and these materials were circulated to selected subject experts for comment. A sample of about 70 documents was used for the purpose.

In developing the materials, we found it did not do justice to our notion of perspective to conceive of our scheme as a simple sum of the major 'isms' commonly referred to in discussions of perspectives. 'Isms' (eg structural-functionalism, Marxism, symbolic interactionism) are characterised in overlapping ways. Writers such as Bernstein[1] and Dawe have drawn out patterns underlying such 'isms'.

Bernstein has identified a range of 'approach paradigms': (a) those who place the emphasis upon the problem of order, as against those who place the emphasis upon the problem of control; (b) those who place the emphasis upon interdependence and dependence, as against those who place the emphasis upon conflict and voluntarism; (c) those who place the emphasis upon how social reality is constructed out of negotiated encounters with others, and those who place the emphasis upon structural relationship; (d) those who emphasise the need to understand the everyday practices of members and the assumptions which make the daily practices work, and those who set up observers' categories and observers' procedures of measurement by means of which they reconstruct the constructions of members.

Even these characterisations tend to take a 'holistic' view of perspectives and, as such, as a basis for the classification of documents, would seem to depend on an overall impression rather than systematic analysis of any given document.

Wallace[2] suggested that it is necessary to employ several dimensions simultaneously in classifying sociological theories. Hoyle[3] adopted a similar procedure in surveying both sociological and non-sociological work concerning education. This approach seemed promising for the systematic analysis of documents in terms of perspective.

The particular dimensions employed by these writers are by no means exhaustive of the range employed currently in the field. To explore the possibility of extending the principle, we selected other dimensions in order to focus upon certain additional perspectives otherwise masked. The materials circulated to subject experts included a synopsis of all these schemas, together with explication, illustration in terms of our sample of documents, and an outline of the practical and conceptual considerations bearing on (a) ultimate selection of the dimensions (and associated categories) to be employed in the system and (b) how these dimensions and categories would be most appropriately related together as a system for the purpose of literature searching.

From a technical point of view, this exercise began to clarify the characteristics we should build into the system. In retrospect, we can see that system construction should in particular take account of the notion of dimension as a defining characteristic of perspective. Whilst in terms of content, the particular dimensions employed in the construction of knowledge may represent a very wide and shifting range, they may be grouped in terms of the aspect of enquiry to which they refer (eg mode of defining subject matter, mode of explanation, mode of verification). If we can define the notion of perspective in these terms, we have a possible framework for the systematic analysis and classification of perspectives in documents which is independent of varying content and definitions of individual perspectives but to which they may readily be related. This type of framework offers a solution to a major problem encountered in the exercise, namely that known perspectives share particular kinds of dimensions in common but, because in terms of content they focus on a different range of conceptual or empirical problems, need to be handled separately in the system in a way which makes the bases of both similarity and difference explicit. It is a framework, too, which allows of response to change in perspectives without disruption of the system.

With regard to appropriateness to conceptualisation in our field, subject experts find on the whole nothing unacceptable in our proposals, with the exception of a spokesman for those who regard systematic information handling per se as interference. Clearly we can do nothing to assist this group. Most agree with our more moderate position, that there is a need for an alternative to systems based, implicitly or explicitly, on objectivist assumptions about knowledge. A reservation was expressed concerning the possible effect on the development of the field in so far as the system might tend to reify the notion of 'approach'. It seems equally valid however to argue that, by making the notion explicit, the system would afford subject experts an opportunity to accelerate the process of synthesising conflicting perspectives, if this is the direction of their interests. Our stress on non-interference is crucial here, together with that on built-in provision for response to change. Another point was that perspectives vary in prestige; hence classification on this basis might be considered invidious. Since such comments were coupled with acceptance of the violence which would be done by perpetuating the myth of objectivism in our system, we do not consider them too damaging. All the arguments of the classic 'equality of opportunity' paradox seem to apply here. Our view is that we have to negotiate a compromise with subject experts which will both avoid serious distortion and minimise labelling effects. There seems to be nothing in the kind of framework we propose which is incompatible with such a compromise. More far-reaching in its implications for our thinking was a point raised by members of our standing working party of subject experts. They questioned that a system which reflected perspectives (whether employed by users or in the literature) as 'products' would be adequate in itself for the purposes of literature searching. For this reason, they recommended, additionally, consideration of perspectives as 'process' (ie as they operate in literature searching) and investigation of the conceptual strategies by which searchers relate to material employing perspectives other than their own. We accept this in principle as an additional factor in system construction requiring investigation.

2. Further thinking

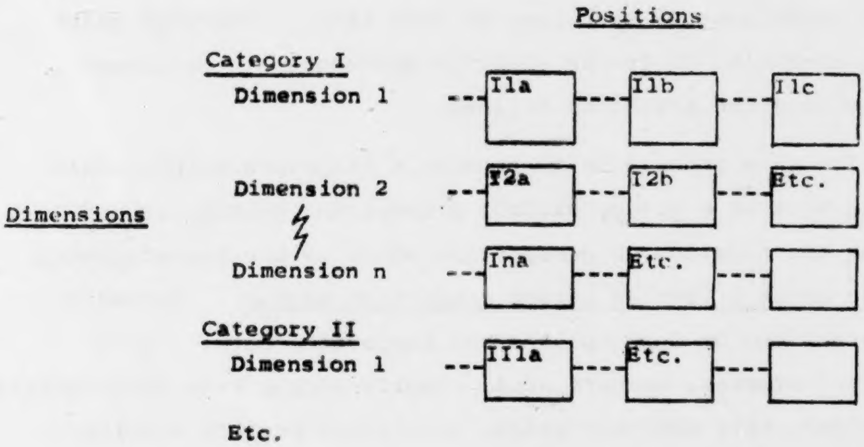
As an alternative procedure for realising perspectives, we explored a form of analytico-synthetic approach. In a planning paper we described the form this might take as follows:

(a) Dimensions analysis

This involves study of theoretical discussion in an article, in which the author characterises his ideas in relation to those of other researchers or of groups of actors, or to general 'taken for granted' opinion. On the basis of this characterisation, documents are plotted on to dimensions.

Nature of dimensions A dimension may comprise two or more alternative positions which may be adopted on a given theoretical issue. Such dimensions cluster into a range of categories. Thus one category might include a dimension on which two possible positions are a biological view of man and a social view of man (theoretical writings contrasting the two may be said to 'belong' to both). The same category might include another dimension which would take account of 'imposed' and 'generated' views of the social world as two possible positions. This category might be described as referring to definitions of the social world.

Presented in diagrammatic form:



Method We are gradually compiling a list of dimensions referred to in the material we are at present working on. These include dimensions in further categories such as level of analysis, 'significant other'; and situation-specific as well as generalised dimensions. Such dimensions, and the 'positions' associated with them, will provide a framework on which we can plot further documents in terms of authors' statements concerning their 'positions', thus creating sets of documents which have some theoretical or conceptual ground in common.

Validation The specific dimensions we identify can be validated only with the authors from whose work they have been derived. Such validation would be combined with validation of our descriptions of their work. The decision to adopt this operational definition of 'dimension' can be validated only at a later stage in terms of effectiveness in an actual context of use.

Relevance in the general context of our work These sets of documents will be a means towards the subsequent characterisation of perspectives.

(b) Characterisation of perspectives

This would involve in-depth study with a group of subject experts.

Problem To understand how subject experts characterise their literature, as a corpus, in terms of perspectives. The group would be asked to identify meaningful combinations of sets of documents. The sets would be created by plotting on dimensions suggested by authors, as described in the previous section. 'Meaningful' combination is defined in terms of expectation that documents might be used together, and hence should be brought together by the system.

We have some experience of exercises of this kind. Thorough piloting would be essential to devise specific procedures to be adopted, but we can outline the method as follows.

Method The main problem is to provide a framework within which the group can work as a group, without prescribing guidelines and hence begging the question of perspective which we are investigating. We propose to adopt a form of paired comparison method. Documents would be represented by descriptions and supporting data. In a series of panel studies, members would compare single sets successively with other single sets and with pairs, triads and further combined sets of documents they identify as 'meaningful'. Each session would also comprise group discussion, in which members would be asked to make explicit, compare and assess the relevance of the bases of their judgments, in the light of their knowledge of the field as a whole. In this way we would hope to arrive at provisional principles for constructing perspectives appropriate to the literature searching situation. The final step in this series of studies would be to investigate appropriate methods of subgrouping material within perspectives, and problems in labelling both perspectives and the topics they define.

Validation The principles we develop would be validated, in a limited way, in terms of their effectiveness in use during the next phase of work, when we simulate the search process, following which they would be subject to revision and reformulation as a set of working principles to guide system construction. Validation with a wider set of opinion would be inappropriate outside an actual context of use. Such validation would follow implementation of the principles in a working system.

Relevance of findings in the general context of our work The principles developed in this phase of our work would be employed in preparing materials for the next phase.

(c) Operational use of perspectives

We are interested here in (1) possible differences between the perspectives by which a corpus of documents may be seen to be structured on the one hand, and the perspectives which users bring to bear on that corpus in searching it on the other. The latter may comprise both perspectives to which an individual is personally and perhaps implicitly committed and perspectives which he adopts explicitly as appropriate to a specific problem under investigation. Any differences we find cannot be assumed to be fixed properties of the situation. This raises (2) the question of interaction amongst perspectives in these categories. In-depth study with subject experts would again be our mode of investigation.

Problem To investigate patterns of change in user perspective during literature searching, as user perspectives either shape responses to the perspectives in the literature or are shaped by them. This would involve comparison of user perspectives (their personal and 'problem' perspectives) before, during (on interaction with each successive perspective in the literature) and after a search.

Method Members would be divided into groups, and each group would be required to agree upon a formulation of a problem for literature searching. Members would then be required to work individually amongst sets of documents (or rather descriptions) constructed in terms of perspectives. The aim would be a simulation of the search situation. Their tools would be a 'map' of perspectives in terms of the dimensions they comprised, and a list of the topics defined by the perspectives. The object would be to locate their problem within a perspective. This would involve a process of gradual approximation unless there were an exactly appropriate system perspective. Each group would be asked to depute a member or members to scan all sets of material, in order to ensure that searching via perspectives had not resulted in relevant documents being overlooked. Group discussion of within and between group differences would be designed to lead to a characterisation of search patterns. Further similar exercises would be carried out over a period to study search patterns over a wider range of search problems.

We may also note in passing that, in the course of our work, discussion with subject experts will bring to light specific assumptions on their part, which we would investigate by small subsidiary studies on an ad hoc basis. The method in these cases would normally be secondary analysis of available data, but small ancillary empirical studies might be incorporated if the need arose, ie if the work could not effectively proceed further until a given point had been clarified.

Relevance of findings in the general context of our work Findings on search patterns would guide decisions as to the ways in which sets of documents, organised in terms of perspectives, would most appropriately be related together as a system.

(d) Specifications for system design

The work described in the previous sections will represent a funnelling down from a large number of sets of documents, defined in terms of 'positions' on dimensions, to a more limited number of combined sets or perspectives. To build these perspectives into a system requires formal principles, which should be guided by the criterion that the generalised search patterns we have identified will not be seriously disrupted by the overall structure we impose.

Nature of specifications Our specifications would comprise principles to guide decisions on the following issues:

i. Selection of system perspectives Our starting point would be our range of 'potentially relevant' perspectives. We would select and build into our system those of value in the process of searching. Specific guidelines would suggest the decisions appropriate to different circumstances. For instance, certain perspectives considered potentially relevant by subject experts might be excluded from the system because they had been found not to relate to search patterns. Additionally, others might be merged because distinctions amongst them were not important in the process of searching.

ii. Interrelationships amongst perspectives Having identified perspectives to be incorporated in the system, we then need to consider the most appropriate way of relating them together. This concerns helpful sequencing of perspectives. Another point is that some perspectives may have been found to serve consistently as decision-making aids, or 'stepping stones' between an initial and a terminal user perspective, rather than as themselves initial or terminal perspectives. In such circumstances, these perspectives would be built into the system in the form of structural links or cross-references across perspectives (eg Perspective A, subset b, see also Perspective C, subset d), not as part of the overall framework of the system.

iii. Substructure within perspectives It would be expected that a perspective would focus on (or be defined by) a distinctive range of substantive or conceptual problems, which would form a structure of subheadings.

iv. Procedures for responding to change The principles we evolve to guide decisions on these issues would be generalised rather than couched in terms of specific current perspectives. As such they would allow changes in the structure of the discipline to be reflected in the system without involving major rethinking of the system. For instance, principles for determining whether and how 'potentially relevant' perspectives are to be built into the system should provide a basis for decisions as to inclusion and treatment of any emergent perspective in the discipline.

Validation Preliminary testing of the principles we formulate would involve their implementation in relation to a sample of documents, and investigation (by correspondence) of each author's opinions of the appropriateness of the context in which, in a future system, his work would be embedded.

In general, however, this would be seen as an exercise in decision-making for the research team, bearing in mind that the principles would be validated in the construction and use of an actual system.

3. Procedures being adopted

It was eventually decided to adopt this general approach. However, it was suggested that we had failed to take account of the possibility that search patterns might be either numerous or complex. If this were the case, we might need to think of users working in post-coordinate fashion ('constructing' perspectives themselves), rather than of a system precoordinated in terms of perspectives. For the purposes of a printed page index, post-coordinate searching tends to be cumbersome, although it is probably to be preferred where on-line computer facilities are available. On the other hand, even a modest degree of precoordination to suit a printed page tool might block certain search patterns. We therefore accepted that questions of format should be left open, along with other questions of physical organisation (eg alphabetical/systematic arrangement), pending clarification of the nature of the search process to be accommodated by the system.

We may therefore characterise the operations involved in the task of system building as follows:

(i) Analysis of further samples of documents, (a) to prepare document substitutes for use in panel studies with subject experts, (b) to develop our prototype scheme of dimensions, and (c) to investigate the range of modes of categorisation associated with each dimension, as evidenced in the literature.

(ii) Classification of document substitutes, in terms of the categories of the scheme, to provide sets of documents to be studied in panel studies. (A given document may be viewed in a number of ways and may form a member of a number of sets.)

(iii) Consultation with subject experts concerning appropriate ways of combining sets of documents, if necessary redefining sets, so as to develop structures consonant with perspectives employed in the field. (Paired comparison exercises, for instance, might provide a starting point for this purpose.)

(iv) Investigation of relative effectiveness, in an operational situation, of organisation by perspective as compared with organisation by category and dimension, leaving users to 'construct' their own perspectives in the latter situation. (This would involve devising a range of alternative search strategies, appropriate structure being that which would be least disruptive of the search process.)

At the present time, we are engaged in (i). The work described by (ii) is dependent on the outcome of our current work. There have been discussions with subject experts concerning the work described in (iii), with which the questions raised in (iv) are interdependent.

III 6B (v) Working papers associated with practical exercises in analysis of documents in terms of dimensions, October 1972 - January 1973.

25.10.72.

Dimensions identified by a brief analysis of a random selection of worksheets [Prepared for meeting on 25.10.72.]

Level of analysis

The most common distinction expressed by this set of authors relates in some sense to level of analysis. Thomas points explicitly to the fact that he has analysed his problem at a macrosociological level and for a clearer understanding of the impact of modernisation analyses at an intra-familial, interpersonal and intermediate level are also necessary. Greenberger stresses she is concentrating on the group rather than the individual level of analysis.

The same kind of consideration, though in a slightly different form, appears to underly decisions concerning the level of explanation which best suits the situation under investigation. The distinction between sociological (in the sense of socio-structural), psychological (in the sense of individual attributes such as personality produced in a social situation) and biological (fixed non-socially produced individual attributes) explanations is particularly common. The distinction is used by some authors to simply group variables for the purposes of analysis. Talking about variables as if they referred to discrete entities (faces of man) often leads on to the further step of explaining the situation in a way consistent with the sets of variables identified. For instance an author handling psychological variables will tend to seek explanations in terms of the individuals unique (to his group) experiences particularly in the process of socialisation. On the other hand an author handling organisational variables might seek to distinguish internal factors such as, in the college situation, social class composition of the students, quality of teaching etc from external factors such as size or prestige (which are externally applied labels on which there is reasonable agreement); he might then proceed to explain the impact of these discrete sets of variables on students by distinguishing effects due to physical and psychological climates.

The following point to distinctions of this kind - Wright, Featherman, Dowse, Hege, Watson, Kamens, Schwartz, Aiken, Sewell, Spillerman, Smith, Crain, Kohn, Richman, Nelson, Eysenck, Little and Lipset.

Further definition of variables

With authors who make distinctions of the previous kind, they often also criticise other researchers for concentrating on one set of variables and ignoring the others hence oversimplifying the situation. On a more restricted scale, at the level of the individual concept, this criticism is also made. Edwards illustrates this point when considering fatherlessness by showing that if one takes the legal definition one ignores important socio-emotional distinctions. Acland stresses the need to view disadvantagement not simply in terms of dilapidated schools but to consider the individual pupil, and to see participation in much more specific terms. Benn stresses the need, not only to consider the extent of comprehensive reorganisation but also its specific type, if adequate predictions are to be made. Little further clarifies the notion of small classes.

The wrong viewpoint

So far the distinctions being made have assumed a certain level on which the authors and the 'other' researchers would agree. In some cases there appears to be no such agreement. The author sees 'other' people as having defined the situation in a completely different way so being unable to see certain factors. Denzin suggests that because childrens activities are viewed as unserious play the serious side, of building up social orders, is ignored. Drucker suggests that if education is looked as a primary (knowledge) industry rather than as a service industry it can be analysed differently. Eysenck stresses the need for a scientific rather than an emotional view of racial differences in IQ.

Linked to this are studies criticising the adoption of the viewpoint of only one partner in a situation (it also relates to further definition). Milner suggests a refocus of attention on immigrants attitudes, rather than hosts attitudes, in studying race relations.

Taken for granted assumptions

Some authors suggest that they are unique because they are considering a problem which others either do not see as a problem or do not see as a studyable problem. Giles falls into the first category by suggesting that speech is normally considered as a static, ie a fixed property of persons, and explores the value of considering it as a dynamic (this could also be seen simply as further definition or wrong viewpoint). Stebbins suggests it is both possible and fruitful to study definition of the situation outside the laboratory.

Discrete theories

None of these authors so far, except Stebbins link themselves to others working in the context of particular theories or viewpoints. Some authors specify that they are using a particular theory, but, because they are not really interested in the viewpoint per se, rarely see it in relation to alternatives. Thomas states he is adopting classical urban theory and Strauss's theory of socialisation without considering what he is not adopting. Bell also does this. Pahl on the other hand does make some attempt to indicate how he sees the interactionist view he adopts as distinct from reference group theory (in terms of unit of consideration ie impact of 'other' or of group) and the possible value of both for studying private and public aspects of behaviour.

Hagstrom suggests he is adopting an input-output view but also considers the impact of other situational factors such as extent of informal communication and discipline orientations.

Conceptions of man

One interesting contrast that emerges is that between the idea that man is a rational animal, or at least acts in accordance with certain fixed attributes he has, or is in some sense irrational and therefore reacts in a non-predictable way.

Studies of the academic selection process illustrate this. Halsey and Trow begin by assuming rationality in mobility - expecting people to move simply on the basis of financial and status rewards offered but find that relative prestige of institutions is a factor. A similar position is taken by Mulkay. Blackburn considers the effect of more information v. the political process itself on decision-making. Lewis considers the effects of academic standards v. other factors in selection.

Related to this are studies which take groups which are related in some objective sense and look for reasons for their obvious different positions. Watson compares the position of women and negroes (both minorities). Goody considers the fallacy of comparing West Africans with other immigrant groups (in the same economic position) for understanding the pattern of fostering. Parsler, in the context of the embourgeoisement thesis, considers the impact of income equalities on life styles of different groups. If income equalities are achieved, are there still 'class' differences?

A particular case of the view that man is or is not a rational animal seems to be at the basis of studies relating people's attitudes and behaviours on the assumption that if rationality could be assumed they would be consistent. Boston considers this in relation to student activism in Italy. Blackburn looks at faculty beliefs and actions in decision making. Inkeles looks at the relationship between attitudes and behaviour.

Policy distinctions

Authors concerned with policy often draw on the conflicting views of actors (normally not other researchers) in given situations in order to address themselves to the issues at hand.

Speaking of libraries Brandon considers the Radcliffe Maud Commission view that libraries are for education in relation to librarians views. Denzin considers views of parents and other countries views of the value of childrens play in relation to observational data. Wolff points to 'commonly held' assumptions about the nature of the child as 'sensitive' or as 'insensitive' to adult emotion. Gretton points to problems associated with the views of graduates and employers that university courses ought to be related to jobs. Kogan illustrates the conflicting views of government and research centres concerning planning of research.

- - - - - - - -

A related set of distinctions concern the appropriate aims of education as seen by different groups. Again these are not usually researchers and are often unspecified.

Taylor asks should the aim of education be equality of opportunity or equality of treatment of LEA's. Oxtoby suggests a conflict between government policy concerning polytechnics and the institutions aims. Maizels contrasts education aimed to fulfil the needs of the labour market and that aimed at equality of opportunity and freedom of choice of the child. Eysenck stresses the need for equality of educational opportunity not a racially biassed system. Benn considers the possible function and aims of education in the 6th form. Miller considers alternatives of education for creativity and for exam success.

With regard to policy oversimplification takes the form of too limited attacks on problems. Sewell for instance suggests need for more focused programs (rather than purely economic) if disadvantagement is to be overcome.

Educational perspectives

The only two items which appear to have identified particular perspectives in education are White and Miller.

Miller distinguishes education suited to the child at various stages of development (ie move to traditional methods at nine when creativity plateau reached) or externally imposed education divisions of policy makers.

White distinguishes radicals and non-radicals in education as being concerned with ends (the kind of society we want) and means respectively.

Spillerman touches on what might be considered an educational perspective (or perhaps a psychological one) by comparing the automatic effects of development on the salience of various rewards and other influences such as social class.

Similarly Wood considers alternative conceptions of the child and of child development as these have guided theories of child rearing.

Methodology

Some methodological considerations are touched on by some of the previous distinctions. For instance oversimplification might be said to have a place in methodology where authors criticise others for using single rather than multiple indicators (or analysing variables separately rather than together) and therefore getting spurious results.

Hamilton suggests multivariate rather than single variable analysis.

Holdaway suggests need for multiple rather than single definitions of variables.

- - - - - - - - -

Oversimplification, due to economic restrictions, might be the cause of authors concerns about the appropriateness of methodology to the situation in hand. Feldman raises the general question concerning the need to adopt appropriate methods to measure college impacts and suggests situations where more complex methods are required. In general the criticism is more specific. Authors may see themselves as distinct because they have adopted a longitudinal rather than a simply cross-sectional approach, or alternatively they may criticise their own work because it is inadequate due to method. Hoge suggests the need for longitudinal study rather than the trend study he did. Holdaway contrasts cross-sectional and longitudinal data. Authors may see themselves as distinct because they have built in added complexity to models of situations to fit the facts better. Sewell sees himself as having done this by further clarifying the concept of social status.

- - - - - - -

The question of levels in methodology probably arises with regard to the authors (Hagstrom, Boruch, Brewer) who compare the merits of comparative and case study data for yielding useful data.

Cause and effect

Related to methodology, but also related to other kinds of distinction, is the idea of cause and automatic effect. Authors often suggest that other researchers have found that A inevitable leads to B but they feel the relationship is more subtle.

Spillerman feels children do not all develop in the same way with regard to the salience of rewards. Inkeles feels modernisation does not, of itself, provide personal disorganisation. Kohn questions whether bureaucracy automatically leads to unthinking, literalistic conformism.

Theory and practice

Two authors drew distinctions between their own aims and those of other researchers in terms of their aims in relation to future research.

Greenberger suggested that the aim was to generate theory rather than to use concepts descriptively. Boruch offers methods of how to implement procedures, ethical codes etc rather than scientific endorsements of codes of ethics to meet major problems or help researchers in specific areas.

Extent of typicality v generalisability of findings

Authors are often concerned that their work is not interpreted too broadly, or alternatively that they have evidence to suggest it can be generalised to other situations.

Hagstrom suggests his work is not adequate for drawing policy implications.

Many authors (Aiken, Kamens, Hagstrom, Kogan) address themselves to the question of the value of British or American data for drawing conclusions about the other country.

Eysenck stresses the need not to generalise too broadly (perhaps a symptom of scientific method) in that evidence collected on American negroes should not be generalised to all negroes.

DW/RAC
25.10.72.

25.10.72.

Questions raised by search for specific dimensions explicitly referred to both directly and indirectly, in current sample of documents. [Prepared for meeting on 25.10.72.]

(1) Most common way is in terms of persons, by way of review of previous research. Quite helpful in indicating relations, but often at level of how one question leads to another, with theoretical framework left implicit. Also degree of selectivity very varied, particularly noticeable in relation to nature of journal (presumably by virtue of audience). Does principle of selectivity also vary? We concluded earlier that organisation by 'significant other' (distinguishing between influence and name-dropping, and between influence in terms of theory, methodology and providing research questions) was worth while. It was not clear whether it would overlap with any other principle of organisation, nor how it would fit in. It is still unclear now. Does not seem unreasonable to think of a given significant other as a 'value' on some dimension, but would be often difficult to say what and why. Perhaps 'influence' is a dimension in itself.

(2) Looking at verbal distinctions, two points emerge. Some people appear to present their work in total isolation from any other. Of the rest who do relate in some way, the immediate impression is that the ways in which they do so have little in common. Even where superficially a similar distinction seems to be made (eg between subjective and objective aspects of a given phenomenon) the point at issue may be a different one (eg the inclusion of social psychological variables as intervening variables in the achievement process / social definitions v resources considered as college output and input variables and the former studied in relation to career allocation).

(3) Seems necessary to consider 'isolationists' and 'relationists' separately. Anything found out about relationists may help with isolates. Problem with isolates is that one cannot tell whether they positively see their work as independent of other work (in which case presumably we must look for some alternative frame of reference), or whether they see the literature in terms of different bodies of work but simply do not feel a need to discuss them in presenting their own. If the latter, the reasons may be extraneous (eg journal policy) or pertinent to us (evaluation implied in ignoring other work). Distinctions made by relationists might suggest what implicit contrasts might be. Inspection of isolationist writing might then draw attention to cues overlooked because of differences in frame of reference.

(4) Relationists' distinctions can provisionally be grouped as follows:

(a) Concern with evidence confirming or contradicting a particular proposition about the behaviour of a given phenomenon (the specific class or instance of the phenomenon may or may not additionally be of interest for its own sake).

Examples are test of embourgeoisement thesis in Australian context, of Blau's theory of interpersonal choice applied to teacher colleague group, theory of developmental sequence in formation of racial attitudes and extent to which US findings hold for UK, theory that bureaucracy makes for conformism in which particular organisations seem to represent a convenience sample.

Writer here seems to identify with a group defined in terms of a particular proposition. Impossible to state whether aim is to 'beat them' or 'join them', because writing up may take advantage of hindsight. Effect is either to consolidate or to undermine what holds the group together. Effect not necessarily dependent on direction of findings. Disconfirmation may suggest refinement of general proposition (eg embourgeoisement thesis exaggerates effects of income) or suggest rejection in favour of some other (bureaucracy and conformism). These are usually presented as suggestions only rather than developed. However, where alternatives are suggested, these perhaps should be considered as linkages to be built into our scheme. If developed the writer would appear to move from one group to another, though the groups are linked by focus on an agreed phenomenon. In this case, is his second location relevant also for our scheme?

In terms of dimension, these people seem to have carved out an agreed phenomenon for study. If we can call this a dimension, then the various propositions about the structure and functioning of this phenomenon, with which they work, might be said to represent alternative values associated with it. The distinctions they make do not seem of themselves to be relevant outside this context, although as a class they may perhaps contrast with others where the aim is something other than confirmation / disconfirmation.

(b) By contrast with the (a) group and its identification (at least for the purposes of the exercise) with a particular proposition, is a group which sets out to evaluate alternative explanations. This may be an artificial distinction in that the (b) group may subsume the (a) group. There is a similar concern with conflicting evidence, and similarly the phenomenon under investigation does not come into question. However, there seems to be a tendency (possibly because more established concepts are used) to label the explanations in more general terms (eg socialisation, sex role differentiation, peer group pressure). These may be referred to in terms of even more general categories (eg psychological v biological - Watson). This may help us in organising the more specific values noted in section (a). There may be a further tendency to difference in that the search is for the theory which offers the most satisfactory explanation rather than seeking confirmation /disproof of the theory in the available evidence.

The comparison of alternatives for best fit pattern obtains also with material in which the focus is on methodology or policy rather than the understanding of a particular phenomenon as such. It may be that in the group of isolates there are specific studies which would find a place in (a) which parallels the theoretical work in its relation with the alternative theories material just discussed. Prescriptive or descriptive writing in which prior assumptions were left unstated would fit this bill.

(c) Representing in their own eyes at least a separate category from (a) are a further group, distinguished by regarding themselves not merely as evaluating other peoples' ideas in an objective way, but as making an original contribution of some kind. They seem to suggest a qualitative difference their work and prior work (eg Featherman contrasts use of a multivariate model with previous research, and includes personality variables as well as social structural variables in his model of the achievement process rather than evaluating individual v social explanations of differences in achievement; Maizels studies conflict between system and personal needs on teacher-pupil relations; Thomas and Weigert refer to a combination of micro and macro-levels in studying industrialisation and conformity; Inkeles looks for an underlying pattern in individual modernisation across cultures; Horowitz and Taylor both add in a temporal dimension in looking at student activism and regional inequality respectively as processes rather than states). Can we say that there is a shift in level, and that the differences in their models are such that sometimes at least they conceive of the phenomenon under study differently, or are even studying a different phenomenon from those whose work they discuss as preceding their own? If so this might provide us with a starting point for grouping phenomena in terms of dimensions (eg a synthesis as in Thomas and Weigert, or abstraction as in Inkeles, is more than the separate elements). Can we distinguish when this is the case from the study in which extra variables are included simply in an attempt to explain a greater proportion of the variance in a given phenomenon. In this latter case the emphasis seems to be as in (b) on the variables in the model than the nature of the model, though the interest is more on the linear relationships amongst specific variables (Sewell) or their relative importance (Little et al on class size and performance) than, as in (b), on alternative classes of variables. (This would fit quite well with CRG theory, although the units we are working with are different.)

There is something of a problem with the type of study in which the omission of a certain range of variables is criticised and the writer then goes on to focus exclusively on the excluded range, presumably leaving others to weave the threads together (eg Kamens on social (extenal) definitions v material (internal) resources in college effects on career allocation). If we take the Atkinson case study seriously there is another problem there in that a manifest / latent distinction is used to contrast the author's view with a widely held view of the selection process, but the conclusion appears to be that however valid it is ill-advised to make the distinction. To the extent that problems such as these recur, handling them is however probably a matter of decision.

Theory and methodology tend to be inseparable here.

(d) Might view (b) as bridging the difference between (a) and (c). In the same way (c) can be viewed as a bridge between (d) and (a + b). Group (c) seen to share the impression of contributing something new rather than consolidating what had gone before. Hence feeling of need to handle differently to mark this somehow. First evidence here of the kind of thing we have had in mind as dimensional distinction. But difference in kind between (c) and (d). The further category (d) seen to represent those whose starting point is different from that of their colleagues. Authors categorised in (c) appear to build on collective work of the individual approaches they dissociate themselves from. Authors considered here present themselves as challenging the assumptions on which available approaches rest and formulate assumptions of their own on which they base their work. Assumptions challenged may be every day or actors' beliefs or values (inferiority of women, meritocratic advancement in academic life, social competence of children and bearing on adult/child relations) or those of observers (Plowden preconceptions, DES prediction of sixth form demands).

Assumptions seem to be challenged at two levels (cf Acland comments on the confounding of variables in Plowden definitions of parental involvement on the one hand, and Denzin's on social evaluations of children's activities as a social order). If this is a distinction we can make reliably, this seems to be another case of a shift in dimension (? taken for granted / problematic), though both concerned with content and thus linked with one subset of (c).

Theory and method tend to be inseparable here.

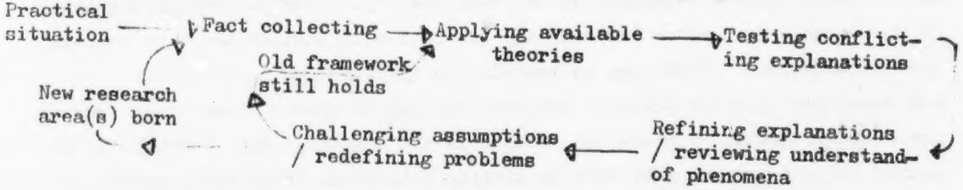
(e) The element of detachment from others in (d) seems to meet up with the consolidation / evaluation element of (b) in a final category, but it is a detachment of non-participation rather than dissociation. This category represents overview of the literature in which, if the writer includes his own work, he treats it (as far as possible) as the work of another. We have few such items in this sample, but one imagines that the picture presented must reflect the state of work in the field at the time of writing and hence could as well point to contrasting explanations to be evaluated (b) as to assumptions which are challenged(d). Detachment / involvement seems unlikely to be either useful or workable as a dimension. Although there might be more of a case for considering whether we could use a distinction between taking the actor's / researcher's standpoint (eg to cater for overviews of methods such as that on studying college impact), methodology still seems better handled as a problem raised by a social situation like any other, and this is more in line with current thinking.

(5) Returning now to the isolates (possibly about $\frac{1}{4}$ of skewed sample?), we can pick out one subset which is cued by starting points such as 'it is obvious, well-known, reasonable to assume etc that ...'. This contrasts with category (d) which shares similar cues. However, in the first case, the writer seeks to draw attention to the consequences, accepting the assumption. In the second, the writer replaces the assumption by another and then reviews the consequences. There are no immediately obvious cues in the rest. We may guess that journal/audience has some bearing on this. Some items take an immediate situation or practical problem as the starting point but this is no reason for not referring to work on similar situations or problems unless we make a distinction between locals / cosmopolitans, as it were. Some are clearly intended to present reliable facts or descriptions simply to make them available as such, and one may guess that there is an implicit assumption that

rigour in data collection assures the impeccability of the data, thus coming under the 'no-one would argue' rubric. However, there is next to no evidence to support this. To the extent that such items extrapolate to trends, or draw attention to consequences in terms of problems or policies, there is a contrast with studies which look to antecedent conditions for explanation and hence prediction. This category requires further thought. It does not seem as intractable as one might have expected. Presumably, at worst, so-called straight description is a type of explanation which is not contrasted with any other, ie (a).

(6) This does not take account of all the individual documents. This does not seem to matter too much at this stage, when we are simply looking for distinctions we might use in starting to build the outline of another framework from the ground-up.

One general point seems worth considering. We have notices in the past that people's distinctions varied considerably in 'level'. Individually, this seems to be a question of where they set their sights, and we have discussed the need to 'round up' their distinctions to arrive at a workable framework. Variable combinations of dimensions is a new possibility. This sample exhibits differences in emphasis which tends to support the variable combinations idea. Looking at the distinctions in terms of levels, however, the 'rounding up' idea seems less attractive, at least as a blanket operation. The question arises whether statement of distinction at a given level reflects a personal view of the relevant literature, or may rather reflect, at the group level, the stage of thinking in a particular area of activity. If this is so we might see as underlying our framework if not explicit in it, a rough model of stages in the life cycle of a research area, eg:


Practical situation
Fact collecting
Applying available theories
Testing conflicting explanations
Old framework still holds
New research area(s) born
Challenging assumptions / redefining problems
Refining explanations / reviewing understand- of phenomena

Variable combinations in this view might be associated with stage in life cycle, in which it seems that certain elements are taken for granted or held constant, whilst others are regarded as problematic and become the focus of investigation. In terms of rounding up, the elements taken for granted are probably considered in very general terms (if indeed at all); those on which investigation is focussed, however, seem to be the subject of comparatively fine distinctions. It seems possible that the nature of the distinctions changes in the course of the cycle from ones of degree (confirmation / disconfirmation of single theory) to ones of kind (when both explanations and models are under scrutiny, and that the range widens as ideas ramify and until a single basic assumption becomes the focus. Where distinctions of kind are made which replace those of degree, however, these tend to be distinctions which have applicability across research areas and not merely within them, involving movement from phenomenon-based to more analytic considerations.

This begins to resurrect some very early ideas, as well as some notions in our theoretical paper about how meaning systems might be expected to behave, though placing them on somewhat different foundations. Is there anything here worth further consideration? For instance, is it fanciful to see not only individuals but research areas 'moving across the board'? Is there any basis here on which we may view simple (or non-existent) perspectives within a single framework along with more complex ones? How do 'isms' relate (we lacked any good ones in our sample)? How far can we be guided any way by authors whose knowledge of an area may be only partial, or may have personal hobby horses?

VW/RAC
25.10.72.

25.10.72.

Notes on discussion of 25.10.72. concerning dimensions analysis

1. Analysis of examples of named 'isms': specific dimensions proposed

   (a) What social reality consists in. It was decided that there were two things here: whether reality is independent of the actor or is what the actor perceives it to be; whether it is pregiven or negotiated in ongoing interaction.

   (b) Logical bases of theories: analytic/synthetic. It was decided that neither this distinction, nor a related (a priori/a posteriori) one could be made to work in our scheme, as far as one could see at present. There were questions of validation and verification to be considered - further thought required.

   (c) Universe of discourse. It was agreed that an individual/social distinction was relevant here, but it seemed possible that level of analysis was only one of several ways of carving up the universe of discourse which we might wish to consider.

   (d) Philosophically normative/explanatory. It was agreed that we needed something like this.

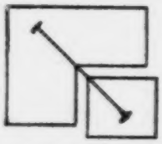
2. Analysis of broader sample of documents

   (a) Despite difference in presentation, it was clear that the two independent analyses of author-indicated dimensions were not seriously incompatible.

   (b) Life-cycle notion was found of interest, though it was noted that, at the genesis of a new research area, the pattern formed by its life cycle might be different from those of current areas. It was seen that it was possible to reduce the ideas down to 4 basic 'attitudes' -

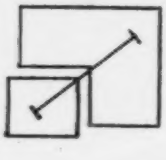
	Treatment of phenomenon	Phenomenon
Conservative	1	3
Initiatory	2	4

It seemed that successively more was taken for granted, and as the focus narrowed, so distinctions became narrow in the range of factors they referred to as well as more specific:



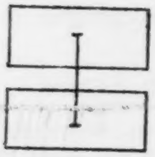
Nature of phenomena questioned.

All other questions subsumed under general dimensions.



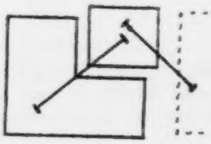
Phenomena not questioned, but nature of 'proper' explanation is

Dimensions tend to refer to general issues but not to such fundamental ones



General approach not in question, but discrepancies in results of applying it attract investigation.

Gives rise to dimensions more at level of (a) model (b) specific research problem.



Concern here with accumulation of detail, filling gaps in agreed framework.

Dimensions again very general if stated at all - relevant dimensions outside this framework or rather between this (research) framework and an alternative (policy) one.

The suggestion was that we might use this as a basis for deciding when differences in the way authors locate themselves should be reflected in the ways we relate them to our eventual scheme. It did not seem likely to form any part of the scheme as such.

3. Content of next phase of work

(a) The problem seemed to be the gulf existing still between dimensions as specified by grand theorists/writers about theory and dimensions as characterised in the 'every day' world of research.

(b) It was decided that the list of general dimensions (see 1 above) required extending as a first task.

(c) Then it would be appropriate to see how far specific dimensions (see 2 above) could be grouped and, at some such intermediate level, might then be viewed appropriately as particular manifestations of a more general dimension. For example 'level of analysis' was clearly an appropriate label for several of the distinctions made by authors in the sample, and this had already been identified as a possible interpretation of universe of discourse.

(d) The object would then be to plot documents on to the resulting 'grid' and on this basis to arrive at a definition of perspective. It was considered that the difficulties of studying the dynamics of perspectives by means of in-depth analysis of clear cut examples of 'isms' were too great for practicability, and hence that this alternative approach was to be preferred.

VW/RAC
10.11.72.

11 & 13.11.72.

Report of 'dimensions' discussion held on 11 & 13 November 1972.

1. A range of theoretical writing had been studied for paired comparisons of different perspectives and 'isms'. Using a set of headings derived from one paper prepared for the previous meeting, the dimensions identified in the survey were built into a rough framework (see attached).

2. Some documents from the sample analysed for the previous meeting were discussed in the light of this framework. A number of these could be accommodated readily within the framework; there were others which it did not touch at any point.

3. This suggested that the immediate problem was one of creating a 'bridge' between dimensions emphasised in documents and dimensions considered important by theoreticians.

4. It had been noted that authors exhibited differing degress of concern with 'fundamental' issues such as the nature of social reality. One may think in terms of 4 broad categories:

(a) Those who appear unconscious of different traditions and seem to have the intention of building on to prior work with no critical appraisal of underlying assumptions.

(b) Those who consider, and perhaps test for 'best fit', alternative models or explanations, but in no way challenge the nature of definitions, or meanings, of the phenomenon under investigation.

(c) Those who identify with one of competing approaches to sociological analysis, which may tend to shape the conceptual model they employ, but do not go so far as to challenge epistemological assumptions.

(d) Those whose starting point is a consideration of their position on epistemological issues, such that their approach and/or model is likely to come under consideration.

5. It was agreed that our scheme must allow us to handle a given document at those levels appropriate to it, but should not force us to make inferences at levels above that of the author's statement of his position. Hence, in terms of our scheme, with the following sections:

I Definition of social reality (DSR)
II Principle of social order (PSO)
III Model and/or nature of unit of study (M/NUS). Possibly also significant other.
IV Specific unit of study (SUS)

Documents in category (a) would be handled in IV.
do. (b) do. III and IV.
do. (c) do. II, III and IV.
do. (d) do. I, II, III, and IV.

It was left an open question whether, if a document was explicitly characterised by its author in terms of sections II and IV but not III, we would make inferences as to its position in terms of III. It was also left open, pro tem, whether a document would be classified on all the dimensions in a given section for which the scheme allowed, or simply those referred to by the author. The problem here is that decisions to work say at the individual/social level of analysis are often taken as self evident and so not referred to explicitly, but are nonetheless relevant across the board.

6. Each combination of dimensions from categories I-III which was found in the documents would in principle form a set of material which would be then grouped by topic (IV). A topic which occurred in several of these sets would be distinctively defined by the dimensions employed in constructing the set to which it belonged. Documents grouped thus in terms of topic could be presented either in dimensional context or alphabetically but tagged for dimensions.

7. This raised a problem in the case of documents in which dimensions of types 1-III were not made explicit. (If it dealt say, with curriculum it might have several possible placings (ie curriculum as variously defined by different combinations of dimensions).) In effect this represents a 'taken for granted' perspective. Basically there are two kinds of document involved: empirical studies such as correlational studies and non-sociological work. The former may for the most part be characterised in terms of the assumptions underlying the statistical analysis carried out, which need not be too controversial. The latter are not in general dimensionless any more than they are perspective-free. The dimensions, however, tend to be social or educational rather than sociological (eg person centred/technology centred, radical/conservative, child oriented/task oriented, subject oriented/pedagogy oriented).

8. It was agreed that these kinds of dimensions should be collected and sorted into an alternative 'grid' comparable to the sociological one. The hope would be eventually to subsume the former under the latter when the scope of these dimensions becomes clearer. Questions concerning purpose of investigation (eg contribution to understanding/guidance for action) would determine whether a document goes through one, other or both 'grids'.

9. A further 'grid', and one which would be expected to remain separate, is a methodological one. Rules for use of this grid, in addition to the main one, for any particular document, need to be devised.

10. It was felt that much remained to be done by way of defining the dimensions noted so far, so as to give splits which are (a) workable, (b) non-controversial, (c) helpful. It was not considered likely that many new dimensions would be added as additional material is processed.

11. Further work on the scheme has been shelved for the immediate term in order to allow us to complete the processing of a new sample of material. Dimensions analysis will accompany description writing for about 60 documents now in the pipe line. The documents will then be split into explicit and implicit perspective sets, C.M. to work on the former and particularly to clarify the 'logic of explanation' dimensions, D.B. and V.W. to work on the latter and to explore the possible nature of a social/educational sub-grid. The plan is to have another meeting in about a fortnight to work on the scheme in the light of handling this sample of documents against the background of the present rough framework and ideas for its development.

Outline of multi-dimensional schemata

A. Theoretical aspects

(1) Definition of social reality

Natural science/social science view of man
Dialectic/non-dialectic
Determinism/voluntarism (specify)
Normative/interpretive
Person/role
Relativism/non-relativism - ethnocentrism
- absolutism
(ie social values do/do not differ
distinguish deep/surface structure)
Epistemological relativism/fixed standards of truth across culture.

(2) Definition of principle of social order

Conflict/consensus
Individualistic/holistic
Alternative views of social structure

(3) Nature of unit of study

Individual/social

Meanings/not - (does author see 'meanings' as problematic?)

Static/dynamic aspects

Cultural relativism/absolutism

Cultural pluralism/monism

System/action

(4) Model and/or specific unit of study

eg (1) Socialisation/external events

eg (2) School as organisation

- - - - -

social system

- - - - -

No. of individuals

(5) Purpose of investigation

eg Contribution to understanding/guidance for action policy (ramify)

B. Methodological aspects

(1) Logic of explanation

In terms of laws - universal/definitional

- statistical

/in terms of particular situation/individual

(2) 'Methodology'

Ideal types

Actor's/observer's viewpoint

Generalisability (extent to which author sees work as generalisable beyond scope of his study)

Form/content

Meanings (eg does author specify particular theory of meanings eg pragmatic)

(3) Methods

eg Observation/questionnaire

VW/RAC
20.11.72.

14-15.12.72.

Report of 'dimensions' discussion, 14-15 December 1972

The particular purpose of this meeting was to discuss the nature of dimensions employed in some educational and other non-sociological literature, and to explore ways of incorporating such dimensions into our classification. This opened up the issue of the nature of the dimensions we have to handle. A number of such dimensions were noted and agreed, eg system/person.

Incorporation in existing scheme

1. There seem to be some grounds for suggesting that such dimensions belong to a set which might be labelled 'values' or 'orientations'. Such a set is missing from our scheme at present, but might be a valuable addition not only for educational material but also for some of our sociological material (on unrest, for example). On the other hand, it was argued that such dimensions were of a different kind from the others so far represented in our scheme, viewed collectively, and cannot validly be regarded as dimensions for the purposes of the scheme.

2. Efforts to specify the basis of this difference, in order to see how 'orientation'-type dimensions might most appropriately be handled, or indeed if they have any place in a scheme of analysis from which perspectives will be identified, have so far proved inconclusive.

3. One possibility considered was that distinctions such as system/person might be regarded as either context-specific statements of more generalised dimensions, or as 'policy' (as contrasted with 'research') equivalents of existing dimensions. In either case, additional dimensions might be handled separately but in parallel with existing ones.

4. An alternative proposal was that the existing scheme be reorganised so that each set of dimensions forms a mirror image of the preceding set, but with the constituent values stated so as successively to limit them either in degree of abstraction or the range of substantive areas to which they might apply. This suggestion might be expected to make provision not only for grand theory but also for middle range theory, and also for 'theorising' (ie analysis of issues) as well as formal theory.

5. Both these suggestions are open to objection on the grounds that they incorporate a 'tracking' or 'streaming' effect, which may prevent the classifier from doing full justice to the material. We know too that subject experts consider this effect undesirable per se. Not only that, our earlier experimentation seemed to show that it is impossible to create an initial sorting frame which represents any acceptable compromise to different groups of users in its practical effects. Apart from this, the distinctions which the classifier is required to make all tend to centre around our original theory/real world one. Not only were the Working Party firmly opposed to this, but, whilst a clear enough distinction can be made in the abstract, it is an exceedingly difficult one to make in terms of actual documents, not least because it tends to require value judgments as to whether theory is 'good' enough to be treated as such.

6. A further possibility suggested as being less open to these criticisms, is one which would seek to give equal opportunity to all documents. This would involve an attempt to define types of dimensions in such a way as to bring together sets of dimensions, irrespective of degree of abstraction, referent, context or order of generality, in terms of underlying notions they may be said to reflect.

7. It seems clear that, as it stands at present, our scheme imposes severe restrictions on the amount of material we can handle in terms of perspectives. Clarification of the intrinsic nature of the dimensions we have found in the literature is plainly an essential next task, from the point of view of knowing what we are handling. As a first step, some educational and psychological material was studied in some detail. (It is obvious that we must reconsider the existing scheme also from the point of view of handling much middle-range sociological writing, including items which build on the work of others rather than taking up and justifying an original theoretical stance for the area concerned.)

Preliminary consideration of nature of our dimensions

1. A particular problem in non-sociological material is that the author does not necessarily locate himself in relation to other researchers. (This was our criterion for recognising valid dimensions in sociological material, but breaks down even with some of this material). An author may instead locate himself, as a researcher, in relation to other groups such as policy makers or to lay conceptions or explanations of a particular phenomenon. This type of work may be contrasted with so-called factual or descriptive writing in a way parallel to the distinction between sociology and sociography.

2. We explored the possibility of working on the basis that, where the author did not explicitly discuss his work in terms of alternative research traditions, it should be required that he adopt a position on the substantive issue in question rather than simply characterise or investigate positions which might be adopted by actors.

3. This requirement, however, gave rise to some unease. In principle it represents a definition of research tradition which is not only narrow but difficult to sustain. For instance, it excludes much behaviourist social psychology, which takes its tradition as self evident.

4. Another possibility suggested was that a distinction be made between analytical dimensions as applied to behaviours or events, and variables representing actors' own categories (eg local/cosmopolitan attitudes v attitudes to specified aspects of education), allowing the former as valid for our purpose. This might give undue weight to correlational and similar studies, although our 'benefit of the doubt' principle would cover this. A more serious question is perhaps the relevance of such dimensions outside a specific class of phenomena (eg persons, or situations in the case of, say, centralisation/decentralisation). This may be one of the factors operating to bring distinctive perspectives to focus on a particular range of empirical or conceptual problems. On the other hand, such studies might be seen to link, say, with conceptualisations of ideal types (such as locals/cosmopolitans), which take the range of application from the individual to the sociological level of analysis, and hence perhaps to a range of perspectives.

5. Earlier discussions with the Working Party tend to suggest that sociologists would welcome the establishment of such links, whereas educationalists would wish to distinguish, so as to avoid, sociological analysis which did not have a direct and obvious bearing on educational situations. Qualification by 'purpose of investigation' was suggested as one way out of this kind of dilemma.

6. Further study of a wider range of material is clearly indicated. One value of this kind of exercise was seen to be the possibility of testing the notion that perspectives do indeed have the effect that theoreticians suggest of focussing on their own distinctive range of topics. Non-sociological material appears to throw up fewer dimensions per document and hence may be convenient for a preliminary test of this assumption. However our discussion of this material so far has served to show that a valid test of the assumption is impossible until we can establish appropriate procedures for characterising the elements of the implicit perspectives we impute to non-sociological

writing, and consider reframing our scheme to take account of the problems raised by this and sociological material alike.

VW/RAC/20.12.72.

January 1973

Progress report on identifying dimensions, January 1973

Earlier work on dimensions had been abortive. We realised that, by drawing on theoretical discussions of perspectives, we had come to define 'dimension' in an unduly restrictive way. The effect was to force us to say that a considerable proportion of our material (sociological as well as educational) would not be classifiable, in any meaningful sense, within our scheme. It was therefore necessary to start again with more open minds as to the ways in which one may conceive of work in the field as being structured.

We had observed several contrasts in the ways in which authors introduced their work:

(1) Gaps, loose ends, discrepancies in substantive findings in their area v. Comments on conceptualisation.

(2) Intention to work within a given framework v. Contrast alternative frameworks.

(3) Intention to identify (build up, extend etc) with a specific area of work v. Dissociation from (rethinking, different approach) a specific area/s of work.

(4) Addressed to bodies of others (eg teachers) distinct from colleagues (eg researchers).

(5) Contrasting the views of bodies of others (eg the lay - 'it is widely assumed that...') with those of colleagues (eg researchers - 'but there is evidence that...').

It seemed possible that prior sorting of material in one, or some combination of, these kinds of ways might produce sets in which dimensions were of a like kind, even though differing across sets. We used a current sample of about 40 documents to test the idea.

There were various operational difficulties in applying these distinctions:

(1) Reliable identification of sets of ideas with categories of people (eg researchers/actors) proved impossible.

(2) Sets of ideas might comprise both the subject matter of, and the framework guiding, an investigation (eg studies of stereotypes).

(3) Identification/dissociation is a subjective question (eg an author may regard his approach as quite new when he adds an extra variable into a model, or may regard himself as building on existing work; in the same way syntheses may or may not be regarded as 'new').

(4) Alternative frameworks may be mentioned merely to dismiss them, or as potentially valid (ie a given amount of space may denote cursory or concise treatment). Again, the framework may be left implicit.

Whilst the distinctions we examined presented difficulties such as these, the exercise identified some specific categories of material which we might have to handle as such (eg material whose starting point is substantive findings, stereotypes). It also suggested further distinctions to be explored:

(1) Questions and answers of interest for own sake v. Interest in implications of questions and answers.

(2) Concern with actual v. ideal situation.

(3) Internal/external view of situation (eg applying a priori framework/taking over a framework from actors, or deriving one from the situation).

(4) Aim to contribute to knowledge or contribute to solution of practical problem.

Such distinctions as these appeared much more viable both operationally and in terms of producing meaningful categories. Flowcharting them, however, produced a very large number of categories. It was decided instead to analyse the superficial structure of the documents, with these kinds of distinctions in mind, to see if we could produce a smaller even if cruder set of categories. For example, contrasting superficial structures found were:

(1) X is important
There are alternative strategies for realising X
Which is better for realising X

(2) Goal of X practices was to do Y
Has there been Y (No)
Agencies A-N have influenced X practices
How can we change A-N, and hence X, to do Y

(3) Want to know about Y
X is a good indicator of Y
What are facts about X
What does this tell us about Y

(4) Propose model X as conceptualisation of phenomenon Y
Derive propositions and test
Is model valid as conceptualisation of Y

Inspection of the analyses of the documents in our sample suggested half a dozen or so recurrent patterns which reflect different styles of research. Our next step was to attempt to specify the questions one would ask to identify each pattern and to flowchart them for decision-making purposes. The major questions appeared to concern the author's stance regarding his subject matter - whether he allowed it to speak for itself, whether he took the actors' perspective(s) or sought to impose a perspective of his own - and his purpose in carrying out a piece of work - eg contribute to understanding for its own sake, assess appropriateness of given goals in given situation etc.

Experimental flowcharting brought us to the conclusion that the crucial factor in our context was author's purpose. A scheme was partially developed on this basis, involving an initial distinction between purpose of 'contributing to understanding'/'guiding action'. We concentrated on types of models employed at this stage. In the first case, we identified models of social/individual structure and functioning, models of relationships amongst properties - internal and external, stereotypes, with a section for facts about a social category, situation etc. In the second, we found various types of ends/means models, eg education as agent of society, responding to societal expectations, reforming society, together with further categories which provided data or ideas from a standpoint considered appropriate in relation to the type of action required. We had also a category for surveys of thinking.

These categories were tested on a new sample of about 100 documents. It was considered to go a good way towards sorting documents appropriately for our purpose, in that the kinds of issues raised in contextualising the author's ideas were of the same kind. However, there was some material which fitted none of our categories of conceptual model, particularly on the 'action' side. Additionally, whilst our categories remained conceptually fairly clear for us, they were not always easy to implement operationally.

It was decided that specific categories of model should be left to emerge, and the team went on to consider how other kinds of dimensions, established in earlier work, might relate to the 'purpose' distinction between 'understanding' and 'action'. These dimensions included methodology, area of enquiry, significant other, together with variables, sample and setting of study. We found that the distinction was not required in any of these cases except methodology. With regard to different categorisations of material, when viewed from the standpoint of these dimensions also, it was decided to leave specific categories to emerge. We therefore adopted, as a prototype framework, the following intellectual organisation. Each cell represents a different standpoint from which the total body of material being handled at any time may be viewed. Associated with each cell are preliminary ideas concerning modes of categorisation, which will provide a basis for discussion with subject experts so as gradually to clarify the extent to which it is possible or desirable to impose further structure, and the form it should take.

It is proposed to continue clarifying definitions of these categories until we have to call a halt in order to illustrate, in our report to OSTI, the point we have reached in building our scheme.

Outline of framework

		Purpose	
	Ways of viewing a corpus of material	Understanding as end in itself	Understanding as required by policy
Whole study	Methodology	(For discussion)	(For discussion)
	Conceptual model	(Level of analysis)	(Means/Ends)
	Significant other	Theory / Research question / Methods	
	Area of enquiry	Theoretical / Substantive / Methodological	
Elements of study	Variables	Clusters	
	Sample	Educational-developmental categories / Other	
	Setting (country)	Alphabetical	

IV General summary, June 1973

This summary complements both our initial chronological overview, and the subsequent groups of papers tracing the development of our thinking in individual areas of our work. We now wish to view our work from a different standpoint, as an integrated programme of investigation, so as to bring into focus an additional range of general points we have so far only touched on in passing.

It must first be noted that there is a central paradox in what we are attempting. In principle, our aim is to devise a non-interfering system, but we recognise that this is a contradiction in terms. Creating a system means imposing structure. The imposition of structure is precisely the task in which the user wishes to engage in searching the literature. To the extent that structure has already been imposed upon the documents, before the user comes to search them, this will interfere with his search. A retrieval system can save him time, but only at the cost of detracting from the quality of the search by distorting it. At the same time there is a practical need, with the increasing volume of research literature, for aids which will make access to relevant literature quicker. In practice, therefore, it is not inappropriate to think in terms of a minimally interfering system. We define this as a system which will interfere only to a degree acceptable to subject experts, and only in ways which will not, in their view, seriously disrupt their searches.

In terms of possible approaches to the task, a major point to make is that a priori structures do not have the necessary goodness of fit with the empirical situation to guide the development of an effective literature searching system. This was our experience in devising procedures for selection and for analysis and description of documents. The point has been driven home on us again more recently. In building our scheme for intellectual organisation, our first idea was that it should reflect perspectivesemployed in the field. Theorists have provided much discussion of the various 'isms' associated with different schools of thought in both sociology and education. However, we could not capture them as entities for our scheme without the danger of reifying them.

Typologies and analyses of the beliefs, assumptions etc implicit in different perspectives suggested that there was a range of common dimensions by which perspectives may be characterised. We experimented with analysis in terms of dimensions and with frameworks constructed in terms of these dimensions. This idea seemed appropriate in that the range of dimensions could be modified over time, without total disruption of the system, to take account of new thinking in the field. The analysis of dimensions was also attractive because of its consonance with the approach of analytical philosophy, which might be expected to give the resulting scheme a substantial measure of objective validity.

Dimensions analysis did not live up to its promise. We soon realised that the philosophical and social-philosophical distinctions discussed by theorists are by no means the whole of the story. We extended our definition of dimension to take account of more middle range work. Once the gap between our ideas and our documents was reduced, the inadequacy of the approach became evident. In the first place it represented too great an oversimplification. It led to a rigidity of structure which prevented us from taking account of the nature and complexity of perspectives in our field. It also carried the danger that, in avoiding the reification of perspectives, we might fall into the trap of reifying dimensions.

Working then independently of preconceived notions of structure, however, we became aware of the many doors we would be closing by particular structures we might impose. We became conscious, for instance in distinguishing between methodology and conceptualisation, and in defining methodology, of a range of distinctions or definitions that we might employ. Even a particular position on a particular dimension (eg individual level of analysis as a mode of conceptualisation) is a complex rather than a unitary notion. There are other points of note. We found that our document descriptions would need to be extended in some respects, for example to relate to classification in terms of methodology, even though authors had judged the descriptions adequate as representations of what they had intended to convey. Certain selection decisions also came into question. We had allowed for the need to modify procedures for selection, and analysis and description of documents, in the light of the developing scheme for intellectual organisation. This simply confirmed our belief that the various processes involved in information processing are highly interdependent. The main point we wish to make here concerns safeguards against interference. We cannot avoid imposing structure, but we must ensure that no major options are overlooked. Additionally, whilst we can include a wide range of options in our system, we do have to select, and selection must be informed.

The approach we have therefore adopted has much in common with grounded theory. We are collecting data from authors, for instance, by means of a simple unstructured enquiry as to the ways in which they characterise their work. We are also inspecting documents for ways in which authors characterise work related to their own. We are allowing categories to emerge from such data. These emerging categories are guiding our thinking on the further development of our procedures for selection, and analysis and description of documents, as well as overall intellectual organisation, and on ways in which we can minimise undue interference. Having said this, we must stress that we are well aware of the limitations of such an approach. It is axiomatic to our work that there be explicit and detailed reasoning guiding all system decisions, and that it be recorded to ensure consistency of future decisions. At the same time, the reasoning must be meaningful, and will differ according to circumstances, if consistency is not to be apparent rather than real. Whilst we are sensitive to some ways of thinking, there are others that we shall not 'see'. We may see structure that is not relevant in users' terms. We feel this approach is justified on the grounds that, for our purpose, the strengths outweigh the weaknesses. We would add to this that the cooperative working relationship we have with subject experts goes a very long way to counteract the weaknesses involved.

It will be clear that our working relationship with subject experts has changed in character since the initiation of our project. Our original plan involved continuing feed-back from a body of subject experts. Individually, each could comment on our work from some but not all of the intellectual perspectives current in the field, but collectively they spanned a very wide range of them. This ability to take different perspectives is of even greater importance in the new relationship which has developed, in which there is a genuine partnership of subject expertise with information science.

Such a partnership is not only appropriate but essential for the development of a non-interfering system. Validation of structure is not a task which can be carried out without the cooperation of subject experts. Their role also extends beyond this. The appropriateness of categories is not simply a question of meaningful representation of the documents as a corpus. The structure must take account of search patterns. Search patterns will reflect the cognitive activity involved in literature searching. Data on this process are hard to come by, and the process clearly cannot be studied 'from the ouside', so reliance on subject experts for guidance in this respect must be heavy. What is happening now is that, as structure is beginning to emerge, so the nature of search processes is becoming clearer. Building the system and devising search patterns are, in other words, two sides of the same coin.

Information science comment on our work has interpreted what we are doing as 'seeking to establish a consensus' amongst subject experts in our field. This is a misunderstanding of our aim which it is most important to correct. Our aim is diametrically opposed to this. Our work is directed towards characterising the differences in thinking amongst different groups in the field, and it is these differences which guide us in developing appropriate procedures for the processing of documents. The only sense in which our system is based on consensus is that groups of subject experts on the whole agree on the ways in which perspectives in the field differ. Consequently we may say that there is a considerable measure of consensus as to the nature of dissensus in thinking in the field. To attempt to eliminate these differences would, of course, be the grossest possible form of interference.

Our aim involves a complete departure from the information science ideal type of a literature searching system as based on consensus. It will not go far enough for many subject experts in that our system will lack sensitivity to nuance. Insensitivity undoubtedly represents interference. On the other hand, apart from the fact that the practicalities of the situation require us to work at the group level, the major argument by which we feel able to justify the imposition of structure of any kind is that we believe there are at the group level, collective perceptions of differences in thinking in the field on which we can draw for guidance. To go about the task with too great a sensitivity on our part would be to tend towards the idiosyncratic, and we would regard this as a more serious form of interference.

Similarly, in terms of responsiveness to developments in the field, the system must change as the empirical situation changes, if it is not to become progressively more interfering, and eventually unusable. However, for the same reason, we propose to work at the level of broad trends rather than individual differences. Responsiveness to change is another notion which is incompatible with the information science ideal type, though the grounds for objection vary from suggesting that change represents an admission of failure to the argument that in economic etc terms it is a practical impossibility. Our view is that even large systems such as MEDLARS, which serves experts in the field of medicine, have had to bow to the realities of the situation. The problem is therefore one of devising the means of managing change that are the most appropriate in all the circumstances of the given situation.

Our greatest problem throughout has been that of ensuring that, in order not to interfere, we had an adequate understanding of the empirical situation of which our system is to form a part. We found that some central assumptions which are taken for granted did not hold when subjected to thorough investigation. We have therefore had to test the ground before us at each step of the way and much of it has crumbled beneath our feet. The interplay between theory and experiment which characterises our work is the means by which we try to think clearly, so as to identify taken for granted or unsupported generalisations which might otherwise go unnoticed, but which should rather be viewed as problematic and requiring investigation. Most of this paper has concentrated on the way theory has guided practice. We conclude with an example of the other aspect of the process, the way in which empirical investigation can lead to extension of theory.

Our investigation of search patterns is not yet far advanced, but it is already producing insights into the nature of the search process which are influencing our thinking. We have long been aware that users are frequently unable to specify the requirements of a search in advance. We had reasoned that, by building perspectives into our system and offering a choice of perspective, we would help the user to orient himself, and thus reduce the time involved in 'funnelling down' to a set of documents appropriate to his needs. This would in no way impede the user wishing to locate documents corresponding to a known set of requirements. We anticipated that by analysing a number of actual searches we might be able to see ways of further assisting the funnelling down operation.

When we began to plan the kind of experimentation that would be required to investigate search patterns, we came to understand that the search process was somewhat more complex than we at first thought. We had viewed the user as selecting from a stable range of alternatives, his choice determined by factors such as subject background, general nature of enquiry and so on. This led us to expect to find a predictable set of search patterns, from which we might select the most widely used to build into our system. Once we added in user perspective as a variable we soon saw that, from the viewpoint of the user, he is confronted by a shifting range of alternatives, since not only will he respond to the documents he encounters, but his perspective on the particular problem in hand is liable to be shaped by them. If so, this means that the direction of a search is rule guided but not, in any sense that is likely to help us in system building, predictable. In other words, although at any single point in a search it is not inappropriate to think of a user as 'having' a perspective, in describing the nature of the search process we must conceptualise the user, in ethnomethodological terms, as 'doing' a perspective. The implications of this view are likely to be far-reaching for understanding of literature searching and for building a system which will not disrupt the dynamics of the process.

V Conclusions

Our research to date is to be summed up less in terms of conclusions arrived at than in terms of issues which we initially regarded as unproblematic and which our work has shown to require investigation. We were advised by subject experts that classification in the sense of 'perception of relationships' was inappropriate for literature searching in our field. We assumed that by a combination of heeding and tempering the conventional wisdom of information science we would, in practice, be able to prove subject experts wrong. Our work may be seen at one level as a vindication of subject experts' opinion on the matter.

It represents more, however, than negation of the appropriateness of conventional techniques. Whilst our experimentation brought us full circle, as it were, we have come not only to accept the viewpoint expressed by subject experts but to reach a deeper understanding of it. This has enabled us to formulate in positive terms an approach to information processing which we believe to be more appropriate than that implicit in conventional techniques.

Basically, information science concludes, because a consensus among subject experts has been found in some fields to provide a basis for effective information service, that such a consensus is a prerequisite for this task in all fields. There is a lack of evidence to support this generalisation. Few would dispute that, in the most general terms, information processing involves the imposition of structure, or that effective information processing will provide appropriate structure, but there is no reason to suggest that a plurality of structures cannot be as effective or even, in some circumstances, more effective than a consensually agreed one. In essence, this means organising material in terms of differences rather than similarities. All the arguments we have been able to muster suggest that this will enhance literature searching rather than impede it in our field.

Once this is accepted, a circularity in information science thinking becomes obvious. Effectiveness is defined as depending on consensus. If consensus need not be assumed, the nature of effectiveness becomes problematic and a new area of investigation is opened up. We have been able to do no more than scratch the surface in the time available to us. Matching, as the mode of search assumed by systems geared to a consensus, will no longer necessarily be preferred. An alternative view, which is much more in keeping with what we know of the ways in which subject experts approach their literature, is one of users interacting with, and responding variously to documents in a corpus, depending on perspective. The effective information service may be defined as one which facilitates the making of such responses by eliminating least essential data, but does not significantly alter the outcome of the interaction. This means that appropriate structure will reflect the major ways in which perspectives differ systematically, rather than translating all material into a single perspective or structure. Such a view thus gives rise to a positive approach to information processing, and indicates both specific issues for exploration and methods for doing so.

In terms of issues for exploration, the concept of perspective has a bearing on all aspects of processing. We may expect differences in perspective to be associated with different definitions of the field. Selection boundaries should take account of such differences, rather than adopting undifferentiated criteria for inclusion and thus opting for a consistency which is apparent rather than real. In the description of individual documents, the basic issue is the nature and amount of information required for 'cuing' responses. This may be expected to vary systematically depending on the nature of the perspective employed in a given document and on user intellectual background factors. Insights into the process of user/document interaction are required so that we may specify a set of procedures which will be sensitive to and supportive of the dynamics of this process without disrupting it.

Fundamental to both selection and description, from which both take their meaning, is the issue of overall organisation of the literature as a corpus. It is in this area that we believe future research effort should be concentrated. If we are correct in thinking that, leaving aside idiosyncratic differences of perspective, cognitive interaction as between subgroups varies systematically rather than randomly, then we may posit a range of cognitive styles which will manifest themselves in strategies of literature searching. The immediate task, then, is one of characterising such cognitive styles as a means towards the end of system building. System building will involve devising ways of organising the literature to ensure equal opportunity (of non-disruption) to a representative range of search strategies.

The notion of interaction has certain implications for methods of investigation. Search strategies will in general be realised in the process of interaction, not formulated *a priori*. Modes of organisation of the literature are appropriate not of themselves but in terms of support for search strategies. Clearly neither can be investigated independently of the other. We envisage therefore a research design in which subject experts and information experts jointly explore these twin problems for system building by engaging in system building along the lines we have described.